

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

% protein - protein search, using sw model

on on: May 21, 2004, 12:34:03 ; Search time 65.8696 Seconds
(without alignments)
64,343 Million cell updates/sec

itle: US-09-980-058-2

effect score: 75

equences: 1 ARAKDELRRTMTM 15

oring table: BLOSUM62

earched: Gapop 10.0 , Gapext 0.5

earched: 1586107 seqs, 282547505 residues

oral number of hits satisfying chosen parameters: 1586107

inimum DB seq length: 0

aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

arabase : A_Geneseq_29Jan04.*

1: Geneseqp1980s.*

2: Geneseqp1990s.*

3: Geneseqp2000s.*

4: Geneseqp2001s.*

5: Geneseqp2002s.*

6: Geneseqp2003as.*

7: Geneseqp2003bs.*

8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

result NO.	Score	Query Match	Length	DB	ID	Description
1	75	100.0	15	4	AAB86600	Human cyt
2	75	100.0	15	4	AAB46264	HCNV IE-1
3	75	100.0	15	4	AAB46134	HCNV type
4	75	100.0	406	2	AAR58703	HCNV IE-e
5	75	100.0	406	2	AAR27275	Human cyt
6	75	100.0	491	4	AAB86544	Human cyt
7	61	81.3	241	6	ABJ38258	Human cyt
8	61	81.3	250	6	ABJ38253	Human cyt
9	57	76.0	15	4	AAB86602	Human cyt
10	56.5	75.3	241	6	ABJ38254	Human cyt
11	56.5	75.3	241	6	ABJ38251	Human cyt
12	56.5	75.3	241	6	ABJ38255	Human cyt
13	53	70.7	15	4	AAB86599	Human cyt
14	52	69.3	10	4	AAB46265	HCNV IE-1
15	52	69.3	10	4	AAB46146	HCNV type
16	52	69.3	10	4	AAB46135	HCNV type
17	50	66.7	241	6	ABJ38252	Human cyt
18	47	62.7	9	4	AAB46272	HCNV IE-1
19	47	62.7	9	4	AAB46142	HCNV type
20	46	61.3	9	4	AAB46271	HCNV IE-1
21	46	61.3	9	4	AAB46141	HCNV type
22	46	61.3	9	4	ABJ38054	Human cyt
23	46	61.3	15	4	AAB46266	HCNV IE-1
24	46	61.3	15	4	AAB46136	HCNV type
25	43	57.3	9	6	ABJ38056	Human cyt

26	43	57.3	241	6	ABJ38259	Human cyt
27	43	57.3	241	6	ABJ38257	Human cyt
28	43	57.3	241	6	ABJ38260	Human cyt
29	43	57.3	241	6	ABJ38256	Human cyt
30	43	57.3	521	6	ABU18101	Protein e
31	42	56.0	80	5	ABP04603	Human ORF
32	42	56.0	141	6	ABU26305	Protein e
33	41	54.7	155	7	ADC00902	Enterohae
34	40	53.3	91	3	AAG54492	Zea mays
35	40	53.3	94	3	AAG54491	Zea mays
36	40	53.3	136	3	AAG54490	Zea mays
37	40	53.3	652	6	ADA54631	Human int
38	40	53.3	652	7	ADD71123	Human int
39	39	52.0	89	3	AAG44849	Zea mays
40	39	52.0	151	3	AAG44847	Zea mays
41	39	52.0	166	5	ABG60005	Human DIT
42	39	52.0	593	2	AAG77757	S-region
43	38.5	51.3	373	7	ADC35636	Wheat SSE
44	38.5	51.3	374	7	ADC35624	Corn SSE1
45	38.5	51.3	374	7	ADC35628	Rice SSE1

ALIGNMENTS

RESULT 1

AAB86600

ID AAB86600 standard; peptide; 15 AA.

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

XX AC AAB86600;

CC lymphocytes, where the stimulated T lymphocytes are expanded and can be
 CC transfused into a patient. AAB86544-AAB86803 represent fragments of the
 CC human cytomegalovirus IE1 and pp65 lower matrix phosphoprotein which are
 CC used to illustrate the method of the invention

XX Sequence 15 AA;

Query Match 100.0%; Score 75; DB 4; Length 15;
 Best Local Similarity 100.0%; Pred. No. 1e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARAKKDELRRKKMMYM 15
 |||||
 DB 1 ARAKKDELRRKKMMYM 15

RESULT 2
 AAB46264
 ID AAB46264 standard; peptide; 15 AA.

XX AC AAB46264;

XX DT 04-APR-2001 (first entry)

XX HCMV IE-1(193-207) antigenic peptide fragment.

XX Antigen, vaccine; human; infection; Immediate-Early 1 Protein; IE-1;
 KW interferon-gamma; CD8+ cell; immunogenic.

XX Human cytomegalovirus.

XX DE19927039-A1.

XX 07-DEC-2000.

XX 04-JUN-1999; 99DE-01027039.

XX 04-JUN-1999; 99DE-01027039.

XX 07-SEP-1999; 99DE-01043702.

XX (KERN/) KERN P.

XX Kern F, Volk H, Reinke P, Paulhaber N, Surel I, Khatamzas E;

XX WPI; 2001-061700/07.

XX New peptides that stimulate cytokine production in T cells, useful for
 PT vaccination against human cytomegalovirus infection and for detecting
 PT immune responses to the virus.

XX Claim 2; Page 7; 10pp; German.

XX This invention describes novel peptides (including Immediate-Early 1
 CC Protein (IE-1) fragments) and their derivatives resulting from deletions,
 CC insertions or substitutions which are capable of inducing the secretion
 CC of interferon-gamma in CD8+ cells. The products of the invention have
 CC immunogenic activity and are useful for the production of vaccines
 CC against human cytomegaloviruses (HCMV) infections and for diagnostics for
 CC the identification of HCMV infections

XX Sequence 15 AA;

Query Match 100.0%; Score 75; DB 4; Length 15;
 Best Local Similarity 100.0%; Pred. No. 1e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARAKKDELRRKKMMYM 15
 |||||
 DB 1 ARAKKDELRRKKMMYM 15

RESULT 3
 AAB46134

ID AAB46134 standard; peptide; 15 AA.

XX AC AAB46134;

XX DT 04-APR-2001 (first entry)

XX HCMV type 16 antigenic peptide SEQ ID NO 2.

XX Antigenic; vaccine; interferon-gamma; tumor necrosis factor-alpha;
 KW CD8+ T cell; HCMV infection; immune response.

XX Human cytomegalovirus.

XX WO200075180-A2.

XX 14-DEC-2000.

XX 02-JUN-2000; 2000WO-DE001854.

XX 04-JUN-1999; 99DE-01027039.

XX 07-SEP-1999; 99DE-01043702.

XX (KERN/) KERN P.

XX Kern F, Volk H, Reinke P, Paulhaber N, Surel I, Khatamzas E;

XX WPI; 2001-061700/07.

XX New peptides that stimulate cytokine production in T cells, useful for
 PT vaccination against human cytomegalovirus infection and for detecting
 PT immune responses to the virus.

XX Claim 1; Page 23; 28pp; German.

XX This invention describes novel peptides (I), their encoding DNA (II) or
 CC their derivatives, that stimulate production of interferon-gamma and
 CC tumor necrosis factor-alpha in CD8+ T cells, particularly those from
 CC humans of suitable HLA type who have been immunized with human
 CC cytomegalovirus (HCMV) (I) and their derivatives are used for
 CC therapeutic or prophylactic vaccination against HCMV infection and as
 CC diagnostic reagents for detecting or quantifying the cellular immune
 CC response to HCMV (from ability to induce production of the specific
 CC cytokines in CD8+ cells), particularly in immune-deficient subjects. Also
 CC DNAs (II) that encode (I), and vectors and plasmids containing (II), are
 CC useful as pharmaceuticals

XX Sequence 15 AA;

Query Match 100.0%; Score 75; DB 4; Length 15;
 Best Local Similarity 100.0%; Pred. No. 1e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARAKKDELRRKKMMYM 15
 |||||
 DB 1 ARAKKDELRRKKMMYM 15

RESULT 4

AAR58703

ID AAR58703 standard; protein; 406 AA.

XX AC AAR58703;

XX 25-MAR-2003 (revised)

XX 29-MAR-1995 (first entry)

XX HCMV IE-exon-4 subunit.

XX HCMV; human cytomegalovirus; major intermediate-early protein; IE gene;
 KW adenovirus; recombinant vaccine.

XX Human cytomegalovirus.

CC lymphocytes, where the stimulated T lymphocytes are expanded and can be
CC transfused into a patient. AAB86544-AAB86803 represent fragments of the
CC human cytomegalovirus IE1 and pp65 lower matrix phosphoprotein which are
CC used to illustrate the method of the invention

XX SQ Sequence 491 AA;

Query Match 100.0%; Score 75; DB 4; Length 491;
Best Local Similarity 100.0%; Pred. No. 0.0003;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARAKDELELRKMMYM 15
DB 193 ARAKDELELRKMMYM 207

RESULT 7
ABJ38258
ID ABJ38258 standard; protein; 241 AA.
XX AC ABJ38258;
XX DT 22-MAY-2003 (first entry)
XX DE Human cytomegalovirus related polypeptide protein SEQ ID No 316.
XX KW Virucide; immunostimulant; cytotoxic T-lymphocyte; CTL; epitope; antigen;
XX KW human cytomegalovirus; HCMV; pp28; pp50; pp65; pp71; pp150; GB; GH; IE-1;
XX KW IE-2; US2; US3; US6; US11; UL18; cytomegalovirus infection; pregnancy;
XX KW transplantation.
XX OS Synthetic.
XX PN WO2003000720-A1.
XX PD 03-JAN-2003.
XX PF 26-JUN-2002; 2002WO-AU000829.
XX PR 26-JUN-2001; 2001AU-00005931.
XX PA (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.
XX PI Khanna R, Elkington RA, Walker SJ;
XX PS WPI; 2003-300379/29.
XX CC New human cytomegalovirus (CMV) cytotoxic T-cell epitope peptide, useful
XX CC for diagnosing, preventing or treating CMV infection, comprises pp28,
XX CC pp50, pp65, pp71, pp150, GB, GH, IE-1, IE-2, US2, US3, US6, US11 or UL18.
XX CC Claim 113; Page 301-302; 308pp; English.

The invention relates to a novel isolated peptide comprising one or more
cytotoxic T-lymphocyte (CTL) epitope(s) of the same or different antigen
of a human cytomegalovirus (HCMV). The HCMV antigens can comprise pp28,
pp50, pp65, pp71, pp150, GB, GH, IE-1, IE-2, US2, US3, US6, US11 or UL18.
The peptide comprises a sequence of about 9-20 contiguous amino acids of
the antigen. The peptide epitopes are useful in diagnosing, preventing or
treating cytomegalovirus infection in humans, and in monitoring immune
responses in various clinical settings (e.g. transplantation or
pregnancy). This sequence represents a synthetic polypeptide protein
relating to the human cytomegalovirus CTL epitopes of the invention

Query Match 81.3%; Score 61; DB 6; Length 241;
Best Local Similarity 82.4%; Pred. No. 0.032;
Matches 14; Conservative 1; Mismatches 0; Indels 2; Gaps 1;

QY 1 ARAKDELELRKMMYM 15
DB 159 ARAKDELELRKMMYM 175

RESULT 8
ABJ38253
ID ABJ38253 standard; protein; 250 AA.
XX AC ABJ38253;
XX DT 22-MAY-2003 (first entry)
XX DE Human cytomegalovirus related polypeptide protein SEQ ID No 311.
XX KW Virucide; immunostimulant; cytotoxic T-lymphocyte; CTL; epitope; antigen;
XX KW human cytomegalovirus; HCMV; pp28; pp50; pp65; pp71; pp150; GB; GH; IE-1;
XX KW IE-2; US2; US3; US6; US11; UL18; cytomegalovirus infection; pregnancy;
XX KW transplantation.
XX OS Synthetic.
XX PN WO2003000720-A1.
XX PD 03-JAN-2003.
XX PF 26-JUN-2002; 2002WO-AU000829.
XX PR 26-JUN-2001; 2001AU-00005931.
XX PA (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.
XX PI Khanna R, Elkington RA, Walker SJ;
XX PS WPI; 2003-300379/29.
XX CC New human cytomegalovirus (CMV) cytotoxic T-cell epitope peptide, useful
XX CC for diagnosing, preventing or treating CMV infection, comprises pp28,
XX CC pp50, pp65, pp71, pp150, GB, GH, IE-1, IE-2, US2, US3, US6, US11 or UL18.
XX CC Claim 113; Page 294-295; 308pp; English.

The invention relates to a novel isolated peptide comprising one or more
cytotoxic T-lymphocyte (CTL) epitope(s) of the same or different antigen
of a human cytomegalovirus (HCMV). The HCMV antigens can comprise pp28,
pp50, pp65, pp71, pp150, GB, GH, IE-1, IE-2, US2, US3, US6, US11 or UL18.
The peptide comprises a sequence of about 9-20 contiguous amino acids of
the antigen. The peptide epitopes are useful in diagnosing, preventing or
treating cytomegalovirus infection in humans, and in monitoring immune
responses in various clinical settings (e.g. transplantation or
pregnancy). This sequence represents a synthetic polypeptide protein
relating to the human cytomegalovirus CTL epitopes of the invention

Query Match 81.3%; Score 61; DB 6; Length 241;
Best Local Similarity 82.4%; Pred. No. 0.032;
Matches 14; Conservative 1; Mismatches 0; Indels 2; Gaps 1;

QY 1 ARAKDELELRKMMYM 15
DB 159 ARAKDELELRKMMYM 175

RESULT 8
ABJ38253
ID ABJ38253 standard; protein; 250 AA.
XX AC ABJ38253;
XX DT 22-MAY-2003 (first entry)
XX DE Human cytomegalovirus related polypeptide protein SEQ ID No 311.
XX KW Virucide; immunostimulant; cytotoxic T-lymphocyte; CTL; epitope; antigen;
XX KW human cytomegalovirus; HCMV; pp28; pp50; pp65; pp71; pp150; GB; GH; IE-1;
XX KW IE-2; US2; US3; US6; US11; UL18; cytomegalovirus infection; pregnancy;
XX KW transplantation.
XX OS Synthetic.
XX PN WO2003000720-A1.
XX PD 03-JAN-2003.
XX PF 26-JUN-2002; 2002WO-AU000829.
XX PR 26-JUN-2001; 2001AU-00005931.
XX PA (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.
XX PI Khanna R, Elkington RA, Walker SJ;
XX PS WPI; 2003-300379/29.
XX CC New human cytomegalovirus (CMV) cytotoxic T-cell epitope peptide, useful
XX CC for diagnosing, preventing or treating CMV infection, comprises pp28,
XX CC pp50, pp65, pp71, pp150, GB, GH, IE-1, IE-2, US2, US3, US6, US11 or UL18.
XX CC Claim 113; Page 294-295; 308pp; English.

The invention relates to a novel isolated peptide comprising one or more
cytotoxic T-lymphocyte (CTL) epitope(s) of the same or different antigen
of a human cytomegalovirus (HCMV). The HCMV antigens can comprise pp28,
pp50, pp65, pp71, pp150, GB, GH, IE-1, IE-2, US2, US3, US6, US11 or UL18.
The peptide comprises a sequence of about 9-20 contiguous amino acids of
the antigen. The peptide epitopes are useful in diagnosing, preventing or
treating cytomegalovirus infection in humans, and in monitoring immune
responses in various clinical settings (e.g. transplantation or
pregnancy). This sequence represents a synthetic polypeptide protein
relating to the human cytomegalovirus CTL epitopes of the invention

Query Match 81.3%; Score 61; DB 6; Length 250;
Best Local Similarity 82.4%; Pred. No. 0.034;
Matches 14; Conservative 1; Mismatches 0; Indels 2; Gaps 1;

QY 1 ARAKDELELRKMMYM 15
DB 195 ARAKDELELRKMMYM 211

RESULT 9
AAB86602
ID AAB86602 standard; peptide; 15 AA.
XX AC AAB86602;
XX DT 20-NOV-2001 (first entry)
XX DE Human cytomegalovirus strain AD169 IE1 peptide fragment SEQ ID 59.
XX KW Antigen-specific stimulation; T-lymphocyte; CD8 stimulation; pp65;
XX KW CD4 stimulation; immuno-stimulation; IE1; lower matrix phosphoprotein.

Query Match 81.3%; Score 61; DB 6; Length 250;
Best Local Similarity 82.4%; Pred. No. 0.034;
Matches 14; Conservative 1; Mismatches 0; Indels 2; Gaps 1;

QY 1 ARAKDELELRKMMYM 15
DB 195 ARAKDELELRKMMYM 211

RESULT 9
AAB86602
ID AAB86602 standard; peptide; 15 AA.
XX AC AAB86602;
XX DT 20-NOV-2001 (first entry)
XX DE Human cytomegalovirus strain AD169 IE1 peptide fragment SEQ ID 59.
XX KW Antigen-specific stimulation; T-lymphocyte; CD8 stimulation; pp65;
XX KW CD4 stimulation; immuno-stimulation; IE1; lower matrix phosphoprotein.

Query Match 81.3%; Score 61; DB 6; Length 250;
Best Local Similarity 82.4%; Pred. No. 0.034;
Matches 14; Conservative 1; Mismatches 0; Indels 2; Gaps 1;

XX Human cytomegalovirus.
XX WO200163286-A2.
XX 30-AUG-2001.
XX 17-FEB-2001; 2001WO-EP001773.
XX 22-FEB-2000; 2000DE-01009341.
XX (KERN/) KERN F.
XX Kern F, Volk H, Reinke P, Faulhaber N, Surel I, Khatamzas B;
XX WPI; 2001-557718/62.
XX Stimulating CD8 or CD4 T cells for the immuno-stimulation of mammals and
XX the determination of the response to the antigen comprises using a
XX synthetic peptide library of the specific antigen.
XX Example; Fig 1; 85pp; German.
XX This sequence represents a novel method for stimulating CD8 or CD4 T
XX cells for the immuno-stimulation of mammals and the determination of the
XX response to an antigen (I). The method comprises (i) dividing the amino
XX acid sequence of the antigen into protein fragments (II) of at least 9
XX amino acids, whereby adjacent or neighbouring fragments are in the whole
XX antigen sequence; (ii) synthesizing a peptide library containing (ii);
XX and (iii) incubating the CD8 and/or CD4 T lymphocytes in a suspension
XX comprising (ii) in a single culture vessel. The method is used to immuno-
XX stimulate T cells of mammals, particularly humans, and is also useful for
XX diagnostic purposes to determine if a mammal, especially human, has
XX previously immunologically responded to a specific protein, and to
XX determine strength of that response. The method is suitable for in vivo
XX or in vitro immuno-stimulation of mammalian, more preferably human T
XX lymphocytes, where the stimulated T lymphocytes are expanded and can be
XX transduced into a patient. AA86544-AA86803 represent fragments of the
XX human cytomegalovirus IE1 and pp65 lower matrix phosphoprotein which are
XX used to illustrate the method of the invention
XX Sequence 15 AA;
XX
XX Query Match 76.0%; Score 57; DB 4; Length 15;
XX Best Local Similarity 100.0%; Pred. NO. 0.01;
XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 2Y 5 KDELRRKKQMYM 15
XX 1 KDELRRKKQMYM 11
XX
XX RESULT 10
XX ABJ38254
XX ID ABJ38254 standard; protein; 241 AA.
XX AC ABJ38254;
XX 22-MAY-2003 (first entry)
XX Human cytomegalovirus related polypeptide protein SEQ ID No 312.
XX Virucide; immunostimulant; cytotoxic T-lymphocyte; CTL; epitope; antigen;
XX human cytomegalovirus; HCMV; pp28; pp50; pp65; pp71; pp150; gB; gH; IE-1;
XX IE-2; US2; US3; US6; US11; UL16; cytomegalovirus infection; pregnancy;
XX transplantation.
XX Synthetic.
XX WO2003000720-A1.
XX 03-JAN-2003.

PF 26-JUN-2002; 2002WO-AU000829.
XX
XX 26-JUN-2001; 2001AU-00005931.
XX (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.
XX Khanna R, Elkington RA, Walker SJ;
XX WPI; 2003-300379/29.
XX
XX New human cytomegalovirus (CMV) cytotoxic T-cell epitope peptide, useful
XX for diagnosing, preventing or treating CMV infection, comprises pp28,
XX pp50, pp65, pp71, pp150, gB, gH, IE-1, IE-2, US2, US3, US6, US11 or UL18.
XX Claim 113; Page 295-296; 308pp; English.
XX
XX The invention relates to a novel isolated peptide comprising one or more
XX cytotoxic T-lymphocyte (CTL) epitope(s) of the same or different antigen
XX of a human cytomegalovirus (HCMV). The HCMV antigens can comprise pp28,
XX pp50, pp65, pp71, pp150, gB, gH, IE-1, IE-2, US2, US3, US6, US11 or UL18.
XX The peptide comprises a sequence of about 9-20 contiguous amino acids of
XX the antigen. The peptide epitopes are useful in diagnosing, preventing or
XX treating cytomegalovirus infection in humans, and in monitoring immune
XX responses in various clinical settings (e.g. transplantation or
XX pregnancy). This sequence represents a synthetic polypeptide protein
XX relating to the human cytomegalovirus CTL epitopes of the invention
XX Sequence 241 AA;
XX
XX Query Match 75.3%; Score 56.5; DB 6; Length 241;
XX Best Local Similarity 53.8%; Pred. No. 0.18;
XX Matches 14; Conservative 1; Mismatches 0; Indels 11; Gaps 1;
XX
XX QY 1 ARAKKD-----ELRRKKQMYM 15
XX 29 ARAKKDELQIKVRDVMVELKQKQMYM 54
XX
XX RESULT 11
XX ABJ38251
XX ID ABJ38251 standard; protein; 241 AA.
XX AC ABJ38251;
XX 22-MAY-2003 (first entry)
XX Human cytomegalovirus related polypeptide protein SEQ ID No 309.
XX Virucide; immunostimulant; cytotoxic T-lymphocyte; CTL; epitope; antigen;
XX human cytomegalovirus; HCMV; pp28; pp50; pp65; pp71; pp150; gB; gH; IE-1;
XX IE-2; US2; US3; US6; US11; UL18; cytomegalovirus infection; pregnancy;
XX transplantation.
XX Synthetic.
XX WO2003000720-A1.
XX 03-JAN-2003.
XX 26-JUN-2002; 2002WO-AU000829.
XX 26-JUN-2001; 2001AU-00005931.
XX (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.
XX Khanna R, Elkington RA, Walker SJ;
XX WPI; 2003-300379/29.
XX
XX New human cytomegalovirus (CMV) cytotoxic T-cell epitope peptide, useful
XX for diagnosing, preventing or treating CMV infection, comprises pp28,
XX pp50, pp65, pp71, pp150, gB, gH, IE-1, IE-2, US2, US3, US6, US11 or UL18.

PS Claim 113; Page 291-292; 308pp; English.

XX The invention relates to a novel isolated peptide comprising one or more
XX cytotoxic T-lymphocyte (CTL) epitope(s) of the same or different antigen
XX of a human cytomegalovirus (HCMV). The HCMV antigens can comprise pp28,
XX pp50, pp65, pp71, pp150, gB, gH, IE-1, IE-2, US2, US3, US6, US11 or UL18.
XX The peptide comprises a sequence of about 9-20 contiguous amino acids of
XX the antigen. The peptide epitopes are useful in diagnosing, preventing or
XX treating cytomegalovirus infection in humans, and in monitoring immune
XX responses in various clinical settings (e.g. transplantation or
XX pregnancy). This sequence represents a synthetic polypeptide protein
XX relating to the human cytomegalovirus CTL epitopes of the invention

XX Sequence 241 AA;

Query Match 75.3%; Score 56.5; DB 6; Length 241;

Best Local Similarity 53.8%; Pred. No. 0.18;

Matches 14; Conservative 1; Mismatches 0; Indels 11; Gaps 1;

2Y 1 ARAKCD-----ELRRKKMYM 15

DB 120 ARAKDELOIKRVDMVELKKRMYM 145

RESULT 12

ABJ38255

ID ABJ38255 standard; protein; 241 AA.

AC ABJ38255;

22-MAY-2003 (first entry)

Human cytomegalovirus related polypeptide protein SEQ ID No 313.

XX Virucide; immunostimulant; cytotoxic T-lymphocyte; CTL; epitope; antigen;
XX human cytomegalovirus; HCMV; pp28; pp50; pp65; pp71; pp150; gB; gH; IE-1;
XX IE-2; US2; US3; US6; US11; UL18; cytomegalovirus infection; pregnancy;
XX transplantation.

XX Synthetic.

WO2003000720-A1.

03-JAN-2003.

26-JUN-2002; 2002WO-AU000829.

26-JUN-2001; 2001AU-00005931.

(COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.

Khanna R, Elkington RA, Walker SJ;

WPI; 2003-300379/29.

XX New human cytomegalovirus (CMV) cytotoxic T-cell epitope peptide, useful
XX for diagnosing, preventing or treating CMV infection, comprises pp28,
XX pp50, pp65, pp71, pp150, gB, gH, IE-1, IE-2, US2, US3, US6, US11 or UL18.

Claim 113; Page 297-298; 308pp; English.

XX The invention relates to a novel isolated peptide comprising one or more
XX cytotoxic T-lymphocyte (CTL) epitope(s) of the same or different antigen
XX of a human cytomegalovirus (HCMV). The HCMV antigens can comprise pp28,
XX pp50, pp65, pp71, pp150, gB, gH, IE-1, IE-2, US2, US3, US6, US11 or UL18.
XX The peptide comprises a sequence of about 9-20 contiguous amino acids of
XX the antigen. The peptide epitopes are useful in diagnosing, preventing or
XX treating cytomegalovirus infection in humans, and in monitoring immune
XX responses in various clinical settings (e.g. transplantation or
XX pregnancy). This sequence represents a synthetic polypeptide protein
XX relating to the human cytomegalovirus CTL epitopes of the invention

XX Sequence 241 AA;

Query Match 75.3%; Score 56.5; DB 6; Length 241;
Best Local Similarity 53.8%; Pred. No. 0.18;

Matches 14; Conservative 1; Mismatches 0; Indels 11; Gaps 1;

QY 1 ARAKCD-----ELRRKKMYM 15

DB 152 ARAKDELOIKRVDMVELKKRMYM 177

RESULT 13

AA86599

ID AA86599 standard; peptide; 15 AA.

XX AA86599;

20-NOV-2001 (first entry)

Human cytomegalovirus strain AD169 IE1 peptide fragment SEQ ID 56.

XX Antigen-specific stimulation; T-lymphocyte; CD8 stimulation; pp65;

XX CD4 stimulation; immuno-stimulation; IE1; lower matrix phosphoprotein.

XX Human cytomegalovirus.

WO200163286-A2.

30-AUG-2001.

17-FEB-2001; 2001WO-EP001773.

22-FEB-2000; 2000DE-01009341.

(KERN/) KERN F.

XX Kern F, Volk H, Reinke P, Faulhaber N, Surel I, Khatazmas E;

WPI; 2001-557718/62.

XX Stimulating CD8 or CD4 T cells for the immuno-stimulation of mammals and
XX the determination of the response to the antigen comprises using a
XX synthetic peptide library of the specific antigen.

XX Example; Fig 1; 85pp; German.

XX This sequence represents a novel method for stimulating CD8 or CD4 T
XX cells for the immuno-stimulation of mammals and the determination of the
XX response to an antigen (I). The method comprises (i) dividing the amino
XX acid sequence of the antigen into protein fragments (II) of at least 9
XX amino acids, whereby adjacent or neighbouring fragments are in the whole
XX antigen sequence; (ii) synthesizing a peptide library containing (II);
XX and (iii) incubating the CD8 and/or CD4 T lymphocytes in a suspension
XX comprising (II) in a single culture vessel. The method is used to immuno-
XX stimulate T cells of mammals, particularly humans, and is also useful for
XX diagnostic purposes to determine if a mammal, especially human, has
XX previously immunologically responded to a specific protein, and to
XX determine strength of that response. The method is suitable for in vivo
XX or in vitro immuno-stimulation of mammalian, more preferably human T
XX lymphocytes, where the stimulated T lymphocytes are expanded and can be
XX transfused into a patient. AA86544-AA86803 represent fragments of the
XX human cytomegalovirus IE1 and pp65 lower matrix phosphoprotein which are
XX used to illustrate the method of the invention

XX Sequence 15 AA;

Query Match 70.7%; Score 53; DB 4; Length 15;

Best Local Similarity 100.0%; Pred. No. 0.047;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARAKDELRKK 11

DB 5 ARAKDELRKK 15

```

RESULT 14
ID AAB46265 standard; peptide; 10 AA.
AC AAB46265;
JT 04-APR-2001 (first entry)
DE HCMV IE-1(198-207) antigenic peptide fragment.
KW Antigen; vaccine; human; infection; Immediate-Early 1 Protein; IE-1;
GW interferon-gamma; CD8+ cell; immunogenic.
XS Human cytomegalovirus.
XX DE19927039-A1.
XX 07-DEC-2000.
XX 04-JUN-1999; 99DE-01027039.
XX 04-JUN-1999; 99DE-01027039.
XX 07-SEP-1999; 99DE-01043702.
XX (KERN/) KERN F.
PI Kern F, Volk H, Reinke P, Faulhaber N, Surel I, Khatamzas E;
PI WPI; 2001-061700/07.
DR New peptides that stimulate cytokine production in T cells, useful for
PT vaccination against human cytomegalovirus infection and for detecting
PT immune responses to the virus.
XX Claim 2; Page 7; 10pp; German.
XX This invention describes novel peptides (including Immediate-Early 1
XX protein (IE-1) fragments) and their derivatives resulting from deletions,
XX insertions or substitutions which are capable of inducing the secretion
XX of interferon-gamma in CD8+ cells. The products of the invention have
XX immunogenic activity and are useful for the production of vaccines
XX against human cytomegaloviruses (HCMV) infections and for diagnostics for
XX the identification of HCMV infections
XX Sequence 10 AA;
XX Query Match 69.3%; Score 52; DB 4; Length 10;
XX Best Local Similarity 100.0%; Pred. No. 0.047;
XX Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
2Y 6 DELRRKKMMYM 15
Db |||||
1 DELRRKKMMYM 10
RESULT 15
ID AAB46146 standard; peptide; 10 AA.
XX AAB46146;
AC AAB46146;
JT 04-APR-2001 (first entry)
DE HCMV type 16 antigenic peptide SEQ ID NO 14.
KW Antigenic; vaccine; interferon-gamma; tumor necrosis factor-alpha;
GW CD8+ T cell; HCMV infection; immune response.
XX Human cytomegalovirus.
XX WO200075180-A2.
XX

```

```

PD 14-DEC-2000.
XX 02-JUN-2000; 2000MO-DE001854.
XX 04-JUN-1999; 99DE-01027039.
PR 07-SEP-1999; 99DE-01043702.
XX (KERN/) KERN F.
XX Kern F, Volk H, Reinke P, Faulhaber N, Surel I, Khatamzas E;
PI WPI; 2001-061700/07.
DR New peptides that stimulate cytokine production in T cells, useful for
PT vaccination against human cytomegalovirus infection and for detecting
PT immune responses to the virus.
XX Claim 1; Page 25; 28pp; German.
XX This invention describes novel peptides (I), their encoding DNA (II) or
XX their derivatives, that stimulate production of interferon-gamma and
XX tumor necrosis factor-alpha in CD8+ T cells, particularly those from
XX humans of suitable HLA type who have been immunized with human
XX cytomegalovirus (hCMV). (I) and their derivatives are used for
XX therapeutic or prophylactic vaccination against hCMV infection and as
XX diagnostic reagents for detecting or quantifying the cellular immune
XX response to hCMV (from ability to induce production of the specific
XX cytokines in CD8+ cells), particularly in immune-deficient subjects. Also
XX DNAs (II) that encode (I), and vectors and plasmids containing (II), are
XX useful as pharmaceuticals
XX Sequence 10 AA;
XX Query Match 69.3%; Score 52; DB 4; Length 10;
XX Best Local Similarity 100.0%; Pred. No. 0.047;
XX Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 6 DELRRKKMMYM 15
Db |||||
1 DELRRKKMMYM 10
Search completed: May 21, 2004, 12:47:26
Job time : 67.8696 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

WM protein - protein search, using sw model

run on: May 21, 2004, 12:45:39 ; Search time 20.8696 Seconds
(without alignments)
37.106 Million cell updates/sec

Title: US-09-980-058-2

Perfect score: 75

Sequence: 1 ARAXXDELRRMMYM 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/2/iaa/PCFUS_COMB.pep.*
- 6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query	ID	Description
1	75	100.0	406	4	US-09-171-699-4
2	75	100.0	406	5	PCT-US94-02107-2
3	42	56.0	93	4	US-09-381-122A-29
4	39	52.0	73	4	US-09-206-551-43
5	39	52.0	585	4	US-09-462-951B-1
6	38	50.7	374	4	US-09-252-991A-31294
7	38	50.7	579	1	US-08-126-564A-31
8	38	50.7	579	5	PCT-US94-09143-31
9	37	49.3	15281	2	US-08-471-119A-2
10	36	48.0	134	4	US-09-732-210-1566
11	36	48.0	159	4	US-09-107-532A-4039
12	36	48.0	165	4	US-09-519-232-58
13	36	48.0	183	4	US-09-489-039A-7793
14	36	48.0	426	4	US-09-252-991A-20262
15	36	48.0	478	4	US-09-252-991A-17502
16	36	48.0	801	3	US-09-104-070-2
17	36	48.0	822	4	US-09-886-319A-64
18	36	48.0	876	1	US-08-717-515-4
19	36	48.0	951	4	US-09-252-991A-26766
20	36	48.0	1084	1	US-08-717-515-6
21	36	48.0	1276	1	US-08-717-515-8
22	35	46.7	148	4	US-09-732-210-1564
23	35	46.7	249	2	US-08-685-992-8
24	35	46.7	249	2	US-09-144-925-8
25	35	46.7	442	3	US-08-821-994-66
26	35	46.7	701	4	US-09-328-352-5717
27	35	46.7	835	4	US-09-758-282B-155

Sequence 243, App
Sequence 259, App
Sequence 263, App
Sequence 12, Appl
Sequence 12, Appl
Sequence 12, Appl
Sequence 12, Appl
Sequence 6, Appl
Sequence 10, Appl
Sequence 4, Appl
Sequence 8, Appl
Sequence 8, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 18, Appl
Sequence 96, Appl
Sequence 590, App
Sequence 38, Appl

ALIGNMENTS

RESULT 1

US-09-171-699-4

; Sequence 4, Application US/09171699

; Patent No. 6448389

; GENERAL INFORMATION:

; APPLICANT: The Wistar Institute of, Anatomy & Biology

; Gonczol, Eva

; Berencsi, Klara

; Kari, Csaba

; TITLE OF INVENTION: No. 6448389el Cytomegalovirus DNA Constructs and

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS: Uses Therefor

; ADDRESS: Howson and Howson

; STREET: Spring House Corporate Center, P.O. Box 457

; CITY: Spring House

; STATE: Pennsylvania

; COUNTRY: USA

; ZIP: 19477

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/171,699

; FILING DATE: 19-Jan-1999

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/015,717

; FILING DATE: 23-APR-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Kodroff, Cathy A.

; REGISTRATION NUMBER: 33,980

; REFERENCE/DOCKET NUMBER: WST666APCT

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 215-540-9200

; TELEFAX: 215-540-5818

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 406 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Query Match 100.0%; Score 75; DB 4; Length 406;

Best Local Similarity 100.0%; Pred. No. 0.00015; Gaps 0;

Matches 15; Conservative 0; Mismatches 0; Indels 0;

PRIOR APPLICATION NUMBER: PCT/IB98/01191
PRIOR FILING DATE: 1998-07-16
PRIOR APPLICATION NUMBER: BP 97112326.0
PRIOR FILING DATE: 1997-07-18
NUMBER OF SEQ ID NOS: 13
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 585
TYPE: PRT
ORGANISM: Mus musculus
S-09-462-951B-1

Query Match 52.0%; Score 39; DB 4; Length 585;
Best Local Similarity 53.8%; Pred. No. 1.1e+02;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
y 1 ARAKDELRRKMM 13
b 319 ARQRKELRRKLL 331

RESULT 6
S-09-252-991A-31294
Sequence 31294, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252.991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 31294
LENGTH: 374
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
S-09-252-991A-31294

Query Match 50.7%; Score 38; DB 4; Length 374;
Best Local Similarity 72.7%; Pred. No. 1e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
2y 1 ARAKDELRRK 11
b 342 ARQRKELRGK 352

RESULT 7
US-08-126-564A-31
Sequence 31, Application US/08126564A
Patent No 5436150
GENERAL INFORMATION:
APPLICANT: Chandrasegaran, Srinivasan
TITLE OF INVENTION: Functional Domains in FokI
TITLE OF INVENTION: Restriction Endonuclease
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cushman, Darby & Cushman
STREET: 1100 New York Ave., N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0,

SOFTWARE: Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/126,564A
FILING DATE: 27-SEPTEMBER-93
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kokulis, Paul N.
REGISTRATION NUMBER: 16,773
REFERENCE/DOCKET NUMBER: PNK/4130/82506/CLB
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861-3503
TELEFAX: 202-822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 579 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-126-564A-31

Query Match 50.7%; Score 38; DB 1; Length 579;
Best Local Similarity 58.3%; Pred. No. 1.5e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
Qy 4 KXDELRRKMM 15
Db 393 KXSELRHKLKV 404

RESULT 8
PCT-US94-09143-31
Sequence 31, Application PC/TUS9409143
GENERAL INFORMATION:
APPLICANT: Chandrasegaran, Srinivasan
TITLE OF INVENTION: Functional Domains in FokI
TITLE OF INVENTION: Restriction Endonuclease
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cushman, Darby & Cushman
STREET: 1100 New York Ave., N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0,
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/09143
FILING DATE: 23-AUG-1994
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/126,564
FILING DATE: 27-SEPTEMBER-93
ATTORNEY/AGENT INFORMATION:
NAME: Kokulis, Paul N.
REGISTRATION NUMBER: 16,773
REFERENCE/DOCKET NUMBER: PNK/4130/82506/CLB
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861-3503
TELEFAX: 202-822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 579 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

```
/ MOLECULE TYPE: peptide
PCT-US94-09143-31
Query Match 50.7%; Score 38; DB 5; Length 579;
Best Local Similarity 58.3%; Pred. No. 1.5e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 AKKDELRKQNM 15
DB 393 KXSELRLHKLKY 404

RESULT 9
US-08-471-119A-2
; Sequence 2, Application US/08471119A
; Patent No. 5827706
; GENERAL INFORMATION:
; APPLICANT: Leitner, Ernst
; APPLICANT: Schneider, Elisabeth
; APPLICANT: Schoergerdorfer, Kurt
; APPLICANT: Weber, Gerhard
; TITLE OF INVENTION: Cyclosporin Synthetase
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSES: No. 5827706artis Corporation
; STREET: 59 Route 10
; CITY: East Hanover
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07936
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471,119A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kassencoff, Melvyn
; REGISTRATION NUMBER: 26,389
; REFERENCE/DOCKET NUMBER: 100-8029/CONT/CONT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 503 8474
; TELEFAX: 201 503 8807
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15281 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Tolypocladium niveum
; STRAIN: ATCC 34921
US-08-471-119A-2
Query Match 49.3%; Score 37; DB 2; Length 15281;
Best Local Similarity 63.6%; Pred. No. 4.1e+03;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 AKKDELRKQNM 13
DB 5214 ASKDRVRQNM 5224

RESULT 10
US-09-732-210-1566
; Sequence 1566, Application US/09732210
; Patent No. 6573361
```

```
/ GENERAL INFORMATION:
; APPLICANT: Bunkers, Greg J.
; APPLICANT: Liang, Jihong
; APPLICANT: Mittanck, Cindy A.
; APPLICANT: Seale, Jeffrey W.
; APPLICANT: Wu, Yonnie S.
; TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use
; FILE REFERENCE: 38-21(15036)B
; CURRENT APPLICATION NUMBER: US/09/732,210
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,513
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: US 60/169,340
; PRIOR FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 1753
; SEQ ID NO 1566
; LENGTH: 134
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-732-210-1566
Query Match 48.0%; Score 36; DB 4; Length 134;
Best Local Similarity 58.3%; Pred. No. 87;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 ARAKDELRKRM 12
DB 32 ANVSKDELEKL 43

RESULT 11
US-09-107-532A-4039
; Sequence 4039, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-8277
; TELEFAX: (781)893-5007
; INFORMATION FOR SEQ ID NO: 4039:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 159 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
```

```
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...159
; SEQUENCE DESCRIPTION: SEQ ID NO: 4039:
JS-09-107-532A-4039
Query Match 48.0%; Score 36; DB 4; Length 159;
Best Local Similarity 58.3%; Pred. No. 1e+02; 2; Indels 0; Gaps 0;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

2Y 3 AKKDELRKMMY 14
   ||| ||| |||
Db 82 AKKAEIRQKXKH 93

RESULT 12
JS-09-519-232-58
; Sequence 58, Application US/09519232
; Patent No. 6528702
; GENERAL INFORMATION:
; APPLICANT: Salmeron, John
; APPLICANT: Weisio, Laura
; APPLICANT: Willits, Michael
; APPLICANT: Mengiste, Tesfaye
; TITLE OF INVENTION: NOVEL PLANT GENES AND USES THEREOF
; FILE REFERENCE: S-30857A/RTP2095
; CURRENT APPLICATION NUMBER: US/09/519,232
; CURRENT FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 58
; LENGTH: 165
; TYPE: PRT
; ORGANISM: Brassica napus
US-09-519-232-58

Query Match 48.0%; Score 36; DB 4; Length 165;
Best Local Similarity 46.2%; Pred. No. 1e+02; 3; Indels 0; Gaps 0;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 AKKDELRKMMY 15
   ||| ||| |||
Db 145 ATPDELRKLLYL 157

RESULT 13
US-09-489-039A-7793
; Sequence 7793, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 7793
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7793

Query Match 48.0%; Score 36; DB 4; Length 183;
Best Local Similarity 53.8%; Pred. No. 1.1e+02; 3; Mismatches 3; Indels 0; Gaps 0;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 ARAKDELRKMM 13
   ||| : |||
```

```
Db 127 AKAMAEAMRRKMM 139

RESULT 14
US-09-252-991A-20262
; Sequence 20262, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20262
; LENGTH: 426
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20262

Query Match 48.0%; Score 36; DB 4; Length 426;
Best Local Similarity 60.0%; Pred. No. 2.4e+02; 4; Mismatches 0; Indels 0; Gaps 0;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 RAKDELRK 11
   ||| : |||
Db 86 RAERDRLRR 95

RESULT 15
US-09-252-991A-17502
; Sequence 17502, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17502
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17502

Query Match 48.0%; Score 36; DB 4; Length 478;
Best Local Similarity 70.0%; Pred. No. 2.7e+02; 2; Mismatches 1; Indels 0; Gaps 0;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 RAKDELRK 11
   ||| : |||
Db 98 RALRDLRKK 107

Search completed: May 21, 2004, 12:50:50
Job time : 21.8696 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

M protein - protein search, using sw model

run on: May 21, 2004, 12:47:34 ; Search time 50.2174 Seconds
(without alignments)
83.314 Million cell updates/sec

title: US-09-980-058-2

effect score: 75

sequence: 1 ARAKXDELRRKMMYM 15

coring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

searched: 1149313 segs, 278921704 residues

total number of hits satisfying chosen parameters: 1149313

inimum DB seq length: 0

maximum DB seq length: 2000000000

post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

database : Published Applications AA:

```
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/RCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	Match	Length	ID	Description
1	75	100.0	406	14	US-10-223-538-4	Sequence 4, Appli
2	43	57.3	521	12	US-10-382-122A-46025	Sequence 46025, A
3	42	56.0	93	12	US-10-382-479-29	Sequence 29, Appl
4	42	56.0	141	12	US-10-282-122A-54229	Sequence 54229, A
5	42	56.0	194	12	US-10-425-114-55920	Sequence 55920, A
6	42	56.0	347	12	US-10-425-114-42246	Sequence 42246, A
7	42	56.0	368	12	US-10-425-114-65610	Sequence 65610, A
8	42	56.0	402	12	US-10-425-114-46457	Sequence 46457, A
9	42	56.0	429	12	US-10-425-114-46776	Sequence 46776, A
10	40	53.3	144	12	US-10-425-114-49225	Sequence 49225, A
11	40	53.3	150	12	US-10-425-114-50553	Sequence 50553, A
12	40	53.3	153	12	US-10-425-114-50553	Sequence 50553, A
13	40	53.3	163	12	US-10-425-114-52657	Sequence 52657, A
14	40	53.3	179	12	US-10-425-114-62266	Sequence 62266, A
15	40	53.3	214	12	US-10-424-599-144412	Sequence 144412,

```
16 40 53.3 221 12 US-10-425-114-38669 Sequence 38669, A
17 40 53.3 304 15 US-10-108-260A-4196 Sequence 4196, Ap
18 40 53.3 437 15 US-10-085-198-162 Sequence 162, App
19 40 53.3 652 15 US-10-094-749-2199 Sequence 2199, Ap
20 39 52.0 73 14 US-10-369-294-43 Sequence 43, Appli
21 39 52.0 585 14 US-10-339-884-1 Sequence 1, Appli
22 38.5 51.3 259 12 US-10-425-114-38384 Sequence 38384, A
23 38.5 51.3 373 14 US-10-268-441-14 Sequence 14, Appli
24 38.5 51.3 374 14 US-10-268-441-6 Sequence 2, Appli
25 38.5 51.3 374 14 US-10-268-441-6 Sequence 6, Appli
26 38 50.7 43 12 US-10-424-599-272870 Sequence 272870,
27 38 50.7 164 12 US-10-424-599-160719 Sequence 160719,
28 38 50.7 682 15 US-10-369-493-22777 Sequence 22777, A
29 38 50.7 828 15 US-10-369-493-901 Sequence 901, App
30 38 50.7 1361 15 US-10-369-493-3209 Sequence 3209, Ap
31 37 49.3 100 12 US-10-424-599-180506 Sequence 180506,
32 37 49.3 120 12 US-10-424-599-190543 Sequence 190543,
33 37 49.3 125 14 US-10-032-201B-90 Sequence 90, Appli
34 37 49.3 133 12 US-10-424-599-194117 Sequence 194117,
35 37 49.3 133 14 US-10-032-201B-80 Sequence 80, Appli
36 37 49.3 135 12 US-10-424-599-201417 Sequence 201417,
37 37 49.3 141 12 US-10-425-114-69427 Sequence 69427, A
38 37 49.3 144 9 US-09-924-358-25 Sequence 25, Appli
39 37 49.3 144 15 US-10-410-764-25 Sequence 25, Appli
40 37 49.3 277 12 US-10-424-599-167982 Sequence 167982,
41 37 49.3 306 10 US-09-800-198-51 Sequence 51, Appli
42 37 49.3 316 12 US-10-425-114-44616 Sequence 44616, A
43 37 49.3 335 10 US-09-934-455-226 Sequence 226, App
44 37 49.3 335 12 US-10-225-068A-834 Sequence 834, App
45 37 49.3 335 15 US-10-374-780A-294 Sequence 294, App
```

ALIGNMENTS

RESULT 1

US-10-223-538-4

; Sequence 4, Application US/10223538

; Publication No. US20030120060A1

; GENERAL INFORMATION:

; APPLICANT: The Wistar Institute of, Anatomy & Biology

; Gonzalez, Eva

; Berencsi, Klara

; Kari, Csaba

; TITLE OF INVENTION: NO. US20030120060A1el Cytomegalovirus DNA Constructs and

; Uses Therefor

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Howson and Howson

; STREET: Spring House Corporate Center, P.O. Box 457

; CITY: Spring House

; STATE: Pennsylvania

; COUNTRY: USA

; ZIP: 19477

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/223,538

; FILING DATE: 19-Aug-2002

; CLASSIFICATION: <unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/09/171,699

; FILING DATE: 19-Jan-1999

; APPLICATION NUMBER: US 60/015,717

; FILING DATE: 23-APR-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Kodroff, Cathy A.

; REGISTRATION NUMBER: 33,980

; REFERENCE/SOCKET NUMBER: NST66APCT

; TELECOMMUNICATION INFORMATION:

TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 406 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-223-538-4

Query Match 100.0%; Score 75; DB 14; Length 406;
Best Local Similarity 100.0%; Pred. No. 0.00034;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARAKDELRRKMYM 15
DB 108 ARAKDELRRKMYM 122

RESULT 2

US-10-282-122A-46025
Sequence 46025, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A

CURRENT APPLICATION NUMBER: US/10/282,122A

CURRENT FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/230,335

PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: 60/230,347

PRIOR FILING DATE: 2000-09-09

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/267,636

PRIOR FILING DATE: 2001-02-09

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 78614

SOFTWARE: PatentIn version 3.1

SEQ ID NO 46025

LENGTH: 521

TYPE: PRT

ORGANISM: Bacillus anthracis

US-10-282-122A-46025

Query Match 57.3%; Score 43; DB 12; Length 521;

Best Local Similarity 57.1%; Pred. No. 72;

Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 ARAKDELRRKMYM 14
DB 460 ARAKDELRRKMYM 473

RESULT 3

US-10-382-479-29

Sequence 29, Application US/10382479

Publication No. US20030224007A1

GENERAL INFORMATION:

APPLICANT: Milligen van, Florine J.

APPLICANT: Bernardus, J.

APPLICANT: Cornelissen, Johannes B.W.J

APPLICANT: Bokhout, Bernard A.

TITLE OF INVENTION: An ex vivo animal or challenge model as method to

measure protective immunity directed against parasites

TITLE OF INVENTION: measure protective immunity directed against parasites

FILE REFERENCE: 2183-4188US

CURRENT APPLICATION NUMBER: US/10/382,479

CURRENT FILING DATE: 2003-03-06

PRIOR APPLICATION NUMBER: US/09/381,122A

PRIOR FILING DATE: 1999-12-23

PRIOR APPLICATION NUMBER: PCT/NL98/00146

PRIOR FILING DATE: 1998-03-11

PRIOR APPLICATION NUMBER: EP 97200730.6

PRIOR FILING DATE: 1997-03-11

NUMBER OF SEQ ID NOS: 32

SOFTWARE: PatentIn Ver. 3.0

SEQ ID NO 29

LENGTH: 93

TYPE: PRT

ORGANISM: Fasciola hepatica

FEATURE:

NAME/KEY: CHAIN

LOCATION: (1)..(93)

OTHER INFORMATION: /Notes="cathepsin L proregion"

US-10-382-479-29

Query Match 56.0%; Score 42; DB 12; Length 93;

Best Local Similarity 70.0%; Pred. No. 19;

Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 6 DELRRKMYM 15

DB 22 DELRRKMYM 31

RESULT 4

US-10-282-122A-54229

Sequence 54229, Application US/10282122A

Publication No. US20040029129A1

GENERAL INFORMATION:

APPLICANT: Wang, Liangsu

APPLICANT: Zamudio, Carlos

APPLICANT: Malone, Cheryl

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Kari

APPLICANT: Zyskind, Judith

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John

APPLICANT: Carr, Grant

APPLICANT: Yamamoto, Robert

APPLICANT: Forsyth, R.

APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A

CURRENT APPLICATION NUMBER: US/10/282,122A

CURRENT FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

```
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 54229
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
US-10-282-122A-54229

Query Match          56.0%; Score 42; DB 12; Length 141;
Best Local Similarity 72.7%; Pred. No. 28;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      2 RAKDELRKRM 12
Db      95 RSKDELRKRM 105

RESULT 5
US-10-425-114-55920
; Sequence 55920, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 55920
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Zea mays subsp. mexicana
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMROTEOSINTE083E12_FLI.pep
US-10-425-114-55920

Query Match          56.0%; Score 42; DB 12; Length 194;
Best Local Similarity 53.8%; Pred. No. 39;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY      2 RAKDELRKRM 14
Db      11 RKQDDMRKVMY 23

RESULT 6
US-10-425-114-42246
; Sequence 42246, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E

; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 55920
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Zea mays subsp. mexicana
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMROTEOSINTE083E12_FLI.pep
US-10-425-114-55920

Query Match          56.0%; Score 42; DB 12; Length 194;
Best Local Similarity 53.8%; Pred. No. 39;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY      2 RAKDELRKRM 14
Db      11 RKQDDMRKVMY 23

RESULT 6
US-10-425-114-42246
; Sequence 42246, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 55920
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Zea mays subsp. mexicana
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMROTEOSINTE083E12_FLI.pep
US-10-425-114-55920

Query Match          56.0%; Score 42; DB 12; Length 347;
Best Local Similarity 53.8%; Pred. No. 70;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY      2 RAKDELRKRM 14
Db      165 RKQDDMRKVMY 177

RESULT 7
US-10-425-114-65610
; Sequence 65610, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 65610
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700152834_FLI.pep
US-10-425-114-65610

Query Match          56.0%; Score 42; DB 12; Length 368;
Best Local Similarity 53.8%; Pred. No. 74;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY      2 RAKDELRKRM 14
Db      185 RKQDDMRKVMY 197

RESULT 8
US-10-425-114-46457
; Sequence 46457, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
```

APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 46457
; LENGTH: 402
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700201552_FLI.pep
US-10-425-114-46457

Query Match 56.0%; Score 42; DB 12; Length 402;
Best Local Similarity 53.8%; Pred. No. 81;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 RAKDELRKMY 14

Db 217 RKQDDMRKMY 229

RESULT 9

US-10-425-114-46776
; Sequence 46776, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:

APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 46776
; LENGTH: 429
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700572930_FLI.pep
US-10-425-114-46776

Query Match 56.0%; Score 42; DB 12; Length 429;
Best Local Similarity 53.8%; Pred. No. 86;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 RAKDELRKMY 14

Db 244 RKQDDMRKMY 256

RESULT 10

US-10-425-114-49225
; Sequence 49225, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:

APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114

; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 49225
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3059-005-H4_FLI.pep
US-10-425-114-49225

Query Match 53.3%; Score 40; DB 12; Length 144;
Best Local Similarity 80.0%; Pred. No. 61;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 AKDELRKMY 12

Db 125 AKDELRKMY 134

RESULT 11

US-10-425-114-65385
; Sequence 65385, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:

APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 65385
; LENGTH: 150
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4762-020-B9_FLI.pep
US-10-425-114-65385

Query Match 53.3%; Score 40; DB 12; Length 150;
Best Local Similarity 80.0%; Pred. No. 64;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 AKDELRKMY 12

Db 132 AKDELRKMY 141

RESULT 12

US-10-425-114-50553
; Sequence 50553, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:

APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 50553
; LENGTH: 153
; TYPE: PRT


```

; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3117-008-D5_FLI.pep
; US-10-425-114-50553

Query Match      53.3%; Score 40; DB 12; Length 153;
Best Local Similarity 80.0%; Pred. No. 65;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      3 AKKDELRKRM 12
DB      134 AKKDELRKI 143

RESULT 13
; US-10-425-114-52657
; Sequence 52657, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 52657
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3068-028-B9_FLI.pep
; US-10-425-114-52657

Query Match      53.3%; Score 40; DB 12; Length 163;
Best Local Similarity 80.0%; Pred. No. 69;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      3 AKKDELRKRM 12
DB      144 AKKDELRKI 153

RESULT 14
; US-10-425-114-62266
; Sequence 62266, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 62266
; LENGTH: 179
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3608-028-E7_FLI.pep
; US-10-425-114-62266

; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3117-008-D5_FLI.pep
; US-10-425-114-50553

Query Match      53.3%; Score 40; DB 12; Length 179;
Best Local Similarity 80.0%; Pred. No. 76;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      3 AKKDELRKRM 12
DB      151 AKKDELRKI 170

RESULT 15
; US-10-424-599-144412
; Sequence 144412, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 144412
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_101417C.1.pep
; US-10-424-599-144412

Query Match      53.3%; Score 40; DB 12; Length 214;
Best Local Similarity 53.8%; Pred. No. 91;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      2 RAKKDELRKRM 14
DB      177 RERGQIRRKRM 189

Search completed: May 21, 2004, 12:52:20
Job time : 51.2174 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

W protein - protein search, using sw model
run on: May 21, 2004, 12:44:59 ; Search time 17.6087 seconds
(without alignments)
81.941 Million cell updates/sec

File: US-09-980-058-2
Perfect score: 75
Sequence: 1 ARAKKDELRKQWYM 15
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : PIR 78.*
1: PIR1.*
2: PIR2.*
3: PIR3.*
4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	75	100.0	491	1 EDBEIC	immediate-early pr
2	75	100.0	491	1 EDBEIS	immediate-early pr
3	42	56.0	141	2 P81426	Ht-transporing tw
4	42	56.0	371	1 S59442	DNA repair protein
5	41	54.7	155	2 D90731	probable endopepti
6	41	54.7	155	2 A85582	probable dipeptid
7	40	53.3	535	2 B95952	trans-regulatory s
8	39	52.0	107	1 VXLJST	hypothetical prote
9	38	50.7	168	2 B6916	probable methionin
10	38	50.7	183	2 E70469	L-serine ammonia-1
11	38	50.7	320	2 B83310	small nuclear ribo
12	38	50.7	462	2 A57120	type II site-speci
13	38	50.7	583	2 A32861	probable ATP-depen
14	38	50.7	607	2 T40906	probable nitrate r
15	38	50.7	828	2 D44990	probable nitrate r
16	38	50.7	828	2 A85860	probable nitrate r
17	38	50.7	828	2 G31015	hypothetical prote
18	38	50.7	1494	2 T26452	probable thiorodox
19	37	49.3	125	2 T50866	thiorodoxin (clone
20	37	49.3	133	2 S58123	conserved coiled-co
21	37	49.3	269	2 T37788	probable hypothet
22	37	49.3	285	2 A22963	folistatin-relate
23	37	49.3	306	2 S38251	hypothetical prote
24	37	49.3	320	2 P98319	hypothetical prote
25	37	49.3	336	2 G44562	probable sugar upt
26	37	49.3	370	2 P96002	integrase (importe
27	37	49.3	419	2 AH0417	coiled coil protei
28	37	49.3	420	2 S43559	heat shock protein
29	37	49.3	463	2 H72365	heat shock protein

30	37	49.3	511	2 S73104	DNA-directed RNA p
31	37	49.3	678	2 A71287	probable cytoplasm
32	37	49.3	800	2 S54623	probable mitochond
33	37	49.3	15281	2 S41309	Cyclosporin synthet
34	36	48.0	78	2 S69531	nolin protein - pu
35	36	48.0	135	2 S48410	ribosomal protein
36	36	48.0	152	2 T31156	hypothetical prote
37	36	48.0	159	2 G82669	ubiquitinone biosynt
38	36	48.0	174	2 D84676	hypothetical prote
39	36	48.0	177	2 T47549	hypothetical prote
40	36	48.0	189	2 A69272	tungsten formylmet
41	36	48.0	217	2 T29388	hypothetical prote
42	36	48.0	221	2 D84780	hypothetical prote
43	36	48.0	263	2 A13081	conserved hypothet
44	36	48.0	263	2 G98204	lactam utilization
45	36	48.0	282	2 D97138	DNA replication pr

ALIGNMENTS

RESULT 1

EDBEIC
immediate-early protein - human cytomegalovirus (strain Towne)
C:Species: human cytomegalovirus, human herpesvirus 5
C>Date: 20-Sep-1984 #sequence_revision 20-Sep-1984 #text_change 16-Jul-1999
C:Accession: A03722; S34420
R:Stenberg, R.M.; Thomsen, D.R.; Stinski, M.F.
J. Virol. 49, 190-199, 1984
A:Title: Structural analysis of the major immediate early gene of human cytomegalovirus
A:Reference number: A03722; UID:84090395; PMID:6317889
A:Accession: A03722
A:Molecule type: DNA
A:Residues: 1-491 <STE>
A:Note: the authors translated the codons GAG, TCC, and CQT for residues 245, 313, and 3
R:Chapman, B.S.; Thayer, R.M.; Vincent, K.A.; Haigwood, N.L.
Nucleic Acids Res. 19, 3979-3986, 1991
A:Title: Effect of intron A from human cytomegalovirus (Towne) immediate-early gene on I
A:Reference number: S34420; UID:91319560; PMID:1650459
A:Accession: S34420
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-48 <CHA>
A:Cross-references: EMBL:M60321; NID:G330624; PIDN:AAA45982.1; FID:G330625
C:Genetics:
A:Introns: 24/2
C:Superfamily: cytomegalovirus immediate-early protein

Query Match 100.0%; Score 75; DB 1; Length 491;
Best Local Similarity 100.0%; Pred. No. 6.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARAKKDELRKQWYM 15
DB 193 ARAKKDELRKQWYM 207

RESULT 2

EDBEIS
immediate-early protein - human cytomegalovirus (strain AD169)
N:Alternate names: UL123 protein
C:Species: human cytomegalovirus, human herpesvirus 5
A:Note: host Homo sapiens (man)
C>Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 16-Jun-2000
C:Accession: S09890
R:Chen, M.S.; Bankier, A.T.; Beck, S.; Bohni, R.; Brown, C.M.; Cerny, R.; Horsnell, T.;
M.: Barrell, B.G.
Curr. Top. Microbiol. Immunol. 154, 125-169, 1990
A:Title: Analysis of the protein-coding content of the sequence of human cytomegalovirus
A:Reference number: S09749; UID:90269039; PMID:2161319
A:Accession: S09890
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA


```

RESULT 7
195952
;Probable dipeptide uptake ABC transporter periplasmic solute-binding protein precursor
;Species: Sinorhizobium meliloti
;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
;Accession: B95952
;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
;oc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
;Title: The complete sequence of the 1.683-kb pSymb megaplasmid from the N2-fixing endo
;Reference number: A95842; MUID:21396508; PMID:11481431
;Accession: B95952
;Status: preliminary
;Molecule type: DNA
;Cross-references: GB:AL591985; PIDN:CAC49282.1; PID:gl5140768; GSPDB:GN00167
;Experimental source: strain 1021, megaplasmid pSymb
;Galibert, P.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, P.; Barloy-Hubler,
;ela, D.; Chain, P.; Cowie, R.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
;L.; Hyman, R.W.; Jones, T.
;Science 293, 668-672, 2001
;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
;ebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
;Reference number: A96039; MUID:21368234; PMID:11474104
;Contents: annotation
;Genetics:
;Gene: orf23; Smb21215
;Genome: plasmid

Query Match 53.3%; Score 40; DB 2; Length 535;
Best Local Similarity 50.0%; Pred. No. 53;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

ZY 1 ARAXKDELRRKQMY 14
|||:||||:
469 ARGELDEAKRKVLY 482

RESULT 8
VKLJST
trans-regulatory splicing protein - human immunodeficiency virus type 2 (isolate ST)
;Alternate names: anti-repression trans-activator; art protein; rev protein; trs protein
;Species: human immunodeficiency virus type 2, HIV-2
;Note: host Homo sapiens (man)
;Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 16-Jul-1999
;Accession: G33943
;Kumar, P.; Hui, H.; Kappes, J.C.; Haggarty, B.S.; Hoxie, J.A.; Arya, S.K.; Shaw, G.M.;
;J. Virol. 64, 890-901, 1990
;Title: Molecular characterization of an attenuated human immunodeficiency virus type 2
;Reference number: A33943; MUID:90112662; PMID:2296086
;Accession: G33943
;Molecule type: genomic RNA
;Residues: 1-107 <KUM>
;Cross-references: GB:M31113; NID:gl339798; PIDN:AAB01357.1; PID:g325754
;Genetics:
;Gene: rev; trs; art
;Introns: 24/1
;Superfamily: AIDS trans-regulatory splicing protein
;Keywords: splicing protein; transcription regulation

Query Match 52.0%; Score 39; DB 1; Length 107;
Best Local Similarity 63.6%; Pred. No. 17;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 RAKDELRRKM 12
|||:||||:
4 RAEELRKL 14

RESULT 9
B96916
hypothetical protein CAC0133 [imported] - Clostridium acetobutylicum
;Species: Clostridium acetobutylicum
;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
;Accession: B96916
;Daly, M.J.; Bennett, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
;J. Bacteriol. 183, 4823-4838, 2001
;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cl
;Reference number: A96900; MUID:21359325; PMID:21359325
;Accession: B96916
;Status: preliminary
;Molecule type: DNA
;Residues: 1-168 <KUR>
;Cross-references: GB:AE001437; PIDN:AAK78117.1; PID:gl5022959; GSPDB:GN00168
;Experimental source: Clostridium acetobutylicum ATCC824
;Genetics:
;Gene: CAC0133

Query Match 50.7%; Score 38; DB 2; Length 168;
Best Local Similarity 70.0%; Pred. No. 39;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 KDELRRKQMY 14
|||:||||:
23 KDELKERKMSY 32

RESULT 10
B70469
probable methionine salvage pathway enzyme E-2/E-2' aq 1975 [similarity] - Aquifex aeoli
;Contains: 5-methylthio-3-oxo-1-penten-1,2-diol dioxygenase (EC 1.13.11.-)
;Species: Aquifex aeolicus
;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 21-Jul-2000
;Accession: B70469
;Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O
;V.
;Nature 392, 353-358, 1998
;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
;Reference number: A70300; MUID:98196666; PMID:9537320
;Accession: B70469
;Status: preliminary; nucleic acid sequence not shown; translation not shown
;Molecule type: DNA
;Residues: 1-183 <AQF>
;Cross-references: GB:AE000766; NID:G2984216; PIDN:AAC07759.1; PID:g2984230; GB:AE00065
;Experimental source: strain VF5
;Genetics:
;Gene: aq 1975
;Superfamily: Klebsiella methionine salvage pathway enzyme E-2/E-2'
;Keywords: metalloprotein; nickel; oxidoreductase
;F;94,95,97,139/Binding site: nickel [Glu, His, His, His] #status predicted

Query Match 50.7%; Score 38; DB 2; Length 183;
Best Local Similarity 70.0%; Pred. No. 42;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 KDELRRKQMY 13
|||:||||:
83 KDELRRKQFL 92

RESULT 11
E83310
L-serine ammonia-lyase (EC 4.3.1.17) [similarity] - Pseudomonas aeruginosa (strain PA01)
;Species: Pseudomonas aeruginosa
;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 21-Jun-2002
;Accession: E83310
;Stover, C.K.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
;Lory, S.; Olson, M.V.
;Nature 406, 959-964, 2000
;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
;Reference number: A82950; MUID:20437337; PMID:10984043
;Accession: E83310

```

```

A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-320 <STO>
A:Cross-references: GB:AE004696; GB:AE0049750; PIDN:AA06071.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA2683
C:Superfamily: threonine dehydratase
C:Keywords: ammonia-lyase; carbon-nitrogen lyase; carbon-oxygen lyase; hydro-lyase

Query Match      50.7%; Score 38; DB 2; Length 320;
Best Local Similarity 72.7%; Pred. No. 71;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ARAKDELRRK 11
DB 288 ARQRKDELRRK 298

RESULT 12
A57120
small nuclear ribonucleoprotein auxiliary factor U2 - mouse
N:Alternate names: Gene Utafl-rs2 protein
C:Species: Mus musculus (house mouse)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
C:Accession: A57120
R:Yamaoka, T.; Hatada, I.; Kitagawa, K.; Wang, X.; Mukai, T.
Genomics 27, 337-340, 1995
A>Title: Cloning and mapping of the U2af1-rs2 gene with a high transmission distortion
A:Reference number: A57120; PMID:9604441; PMID:7558001
A:Accession: A57120
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-462 <RES>
A:Cross-references: GB:D45205; NID:9927658; PIDN:BA08143.1; PID:9927659

Query Match      50.7%; Score 38; DB 2; Length 462;
Best Local Similarity 58.3%; Pred. No. 1e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 ARAKDELRRKM 12
DB 100 ARKRDELRRKL 111

RESULT 13
A32861
type II site-specific deoxyribonuclease (EC 3.1.21.4) - Flavobacterium sp.
N:Alternate names: type II restriction enzyme
C:Species: Flavobacterium sp.
C>Date: 22-Nov-1989 #sequence_revision 22-Nov-1989 #text_change 08-Oct-1999
C:Accession: A32861; JQ0034
R:Kita, K.; Kotani, H.; Sugisaki, H.; Takanami, M.
J. Biol. Chem. 264, 5751-5756, 1989
A>Title: The FokI restriction-modification system. Organization and nucleotide sequences
A:Reference number: A32861; PMID:89174630; PMID:2784436
A:Accession: A32861
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-593 <KIT>
A:Cross-references: GB:J04623; NID:9148705; PIDN:AA24927.1; PID:9148707
R:Looney, M.C.; Moran, L.S.; Jack, W.E.; Feehery, G.R.; Benner, J.S.; Slatko, B.E.; Wills
Gene 80, 193-208, 1989
A>Title: Nucleotide sequence of the FokI restriction-modification system: separate strai
A:Reference number: A91615; PMID:90060769; PMID:2684765
A:Accession: JQ0034
A:Molecule type: DNA
A:Residues: 5-583 <LOO>
C:Comment: This enzyme is one of the two enzymes of FokI, a type-IIS restriction-modific
C:Comment: This enzyme recognizes the asymmetric sequence 5'-GGATG/3'-CCTAC.
C:Genetics:
A:Gene: fokIR
C:Keywords: hydrolase; restriction modification system

```

```

Query Match      50.7%; Score 38; DB 2; Length 583;
Best Local Similarity 58.3%; Pred. No. 1.3e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 KXDELRRKMYM 15
DB 397 KXSELRRHLKYV 408

RESULT 14
T40906
probable ATP-dependent DNA helicase - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 24-Nov-2003
C:Accession: T40906
R:Murphy, L.; Harris, D.; Lyne, M.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, December 1998
A:Reference number: Z21956
A:Accession: T40906
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-607 <MUR>
A:Cross-references: EMBL:AL034490; PIDN:CAA22471.1; GSPDB:GN00068; SPDB:SPCC136.02c
A:Experimental source: strain 972h-; cosmid c126
C:Genetics:
A:Gene: SPDB:SPCC136.02c
A:Map position: 3
A:Introns: 11/3; 72/3; 105/1; 501/3; 571/3
C:Superfamily: Ku70 antigen

Query Match      50.7%; Score 38; DB 2; Length 607;
Best Local Similarity 63.6%; Pred. No. 1.3e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 KXDELRRKMYM 14
DB 294 KXDELRRSYSY 304

RESULT 15
D64990
probable nitrate reductase (EC 1.7.99.4) - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C>Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 23-Sep-2002
C:Accession: D64990
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A>Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; PMID:97426617; PMID:9278503
A:Accession: D64990
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-828 <BLAT>
A:Cross-references: GB:AE000309; GB:U00096; NID:91788520; PIDN:AA05266.1; PID:91788534.
A:Experimental source: strain K-12, substrain MGI655
C:Genetics:
A:Gene: napA
C:Superfamily: Alkaligenes eutrophus nitrate reductase A chain
C:Keywords: oxidoreductase

Query Match      50.7%; Score 38; DB 2; Length 828;
Best Local Similarity 66.7%; Pred. No. 1.7e+02;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 AKXDELRRKMYM 14
DB 581 AKXPELRGKLY 592

Search completed: May 21, 2004, 12:50:06
Job time : 19.6087 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

M protein - protein search, using sw model

Run on: May 21, 2004, 12:34:48 ; Search time 13.6957 Seconds
(without alignments)
57.029 Million cell updates/sec

Title: US-09-980-058-2

Perfect score: 75

Sequence: 1 ARAKKDELRRKMMYM 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	75	100.0	491	1 VIE1_HCMVA	P13202 human cytom
2	75	100.0	491	1 VIE1_HCMVT	P03169 human cytom
3	42	56.0	371	1 RA14_YEAST	P28519 saccharomyc
4	39	52.0	107	1 REV_HV2ST	P20870 human immun
5	38	50.7	325	1 SWK_FUSNN	Q89g3 fusobacteri
6	38	50.7	462	1 U2R2_MOUSE	Q82377 mus musculu
7	38	50.7	583	1 T2F1_FLAOK	P14870 Flavobacter
8	38	50.7	828	1 NAPA_ECOLI	P13937 escherichia
9	37	49.3	125	1 THIH_PICMA	Q65049 picea maria
10	37	49.3	133	1 THH2_ARATH	Q38879 arabidopsis
11	37	49.3	245	1 D3E2_BUCBP	Q89an3 buchnera ap
12	37	49.3	269	1 Y2R8_SCHPO	O13743 schizosacch
13	37	49.3	306	1 PSL1_MOUSE	Q82356 mus musculu
14	37	49.3	463	1 HSLU_THERMA	Q9wy22 thermotoga
15	37	49.3	677	1 CFPA_TREPA	Q56336 treponema p
16	37	49.3	800	1 PT27_YEAST	P32606 saccharomyc
17	36	48.0	78	1 HOL1_BHRP1	P51727 bacteriopho
18	36	48.0	134	1 RS24_YEAST	P26782 saccharomyc
19	36	48.0	255	1 YGH7_AGRF5	Q8u822 agrobacteri
20	36	48.0	306	1 FSL1_RAT	Q82632 rattus norv
21	36	48.0	354	1 GB14_XENLA	O73819 xenopus lae
22	36	48.0	426	1 H1SX_HELMO	Q8gdp4 heliobacilli
23	36	48.0	430	1 T3K2_MOUSE	Q92297 mus musculu
24	36	48.0	474	1 ALGX_PSEAE	Q51372 pseudomonas
25	36	48.0	546	1 SYE_CLOPS	Q8xlp3 clostridium
26	36	48.0	822	1 PER_HUMAN	P16581 homo sapien
27	36	48.0	1360	1 CING_XENLA	Q9pdc7 xenopus lae
28	35	46.7	69	1 YDH2_ECOLI	P77274 escherichia
29	35	46.7	102	1 HVRA_RHOCA	P42505 rhodobacter
30	35	46.7	148	1 RS24_RHIRA	P14249 rhizomucor
31	35	46.7	160	1 YMT0_YEAST	Q04210 saccharomyc
32	35	46.7	277	1 YK33_YEAST	P36103 saccharomyc
33	35	46.7	283	1 RPAZ_ECOLI	P27241 escherichia

Q00123 ictalurid h
Q26110 methanobact
Q28294 canis famil
P23625 drosophila
P91950 homarus ame
P50148 homo sapien
P21279 mus musculu
P22471 rattus norv
P38410 xenopus lae
P38408 bos taurus
Q95837 homo sapien
P30677 mus musculu

ALIGNMENTS

RESULT 1
VIE1_HCMVA
ID VIE1_HCMVA STANDARD; PRT; 491 AA.
AC P13202;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-JUL-1993 (Rel. 26, Last annotation update)
DE 55 kDa immediate-early protein 1 (IE1).
GN UL123.
OS Human cytomegalovirus (strain AD169).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10360;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85194814; PubMed=2986374;
RA Akrigg A., Wilkinson G.W.G., Oram J.D.;
RT "The structure of the major immediate early gene of human
cytomegalovirus strain AD169."
RL Virus Res. 2:1107-121(1985).
RN [2]
RP COMPLETE GENOME.
RX MEDLINE=90269039; PubMed=2161319;
RA Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,
Horsnell T., Hutchinson C.A. III, Kourazides T., Martignetti J.A.,
Predie E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;
RT "Analysis of the protein-coding content of the sequence of human
cytomegalovirus strain AD169."
RL Curr. Top. Microbiol. Immunol. 154:125-169(1990).
CC -!- FUNCTION: THE ACTIVATION OF THE E1.7 PROMOTER BY E1.2 IS
AUGMENTED BY THE IE1 PROTEIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL; X17403; CAA35325.1; -;
DR EMBL; M21295; AAA45980.1; -;
PIR; S09890; EDBEM5.
KW Early protein; Phosphorylation; Nuclear protein.
SQ SEQUENCE 491 AA; 55109 MW; CC5966B00CD8C8B4 CRC64;

Query Match 100.0%; Score 75; DB 1; Length 491;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARAKKDELRRKMMYM 15
|||
DB 193 ARAKKDELRRKMMYM 207

RESULT 2

```

VIE1_HCMVT
ID VIE1_HCMVT STANDARD; PRT; 491 AA.
AC P03169.
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 55 kDa immediate-early protein 1 (IE1).
EN UL123.
OS Human cytomegalovirus (strain Towne).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
CX NCBI_TaxID=10363;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84090395; PubMed=63117889;
RA Stenberg R.M., Thomsen D.R., Stinski M.F.;
RT "Structural analysis of the major immediate early gene of human
RT cytomegalovirus."
RL J. Virol. 49:190-199(1984).
CC -!- FUNCTION: THE ACTIVATION OF THE IE1.7 PROMOTER BY IE2 IS
CC AUGMENTED BY THE IE1 PROTEIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M11630; AAA45979.1;
DR EMBL; M11628; AAA45979.1; JOINED.
DR EMBL; M11629; AAA45979.1; JOINED.
DR F03722; EDBEIC.
KW Early protein; Phosphorylation; Nuclear protein.
FT CONFLICT 96 96 V -> L (IN REF. 1; AAA45979).
SQ SEQUENCE 491 AA; 55178 MW; E6041928A91A5867 CRC64;

Query Match 100.0%; Score 75; DB 1; Length 491;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ARAKKDELRRKKWYM 15
Db 193 ARAKKDELRRKKWYM 207

RESULT 3
RA14_YEAST STANDARD; PRT; 371 AA.
AC P28519;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE DNA repair protein RAD14.
GN RAD14 OR YMR201C OR YMR325.02C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A. AND MUTANT RAD14-2.
RX MEDLINE=97197968; PubMed=9046084;
RA Jones G.W., Reed S.H., Waters R.;
RT "Characterization of the rad14-2 mutant of Saccharomyces cerevisiae:
RT implications for the recognition of UV photoproducts by the Rad14
RT protein."
RL Yeast 13:31-36(1997).
RN [2]
RP SEQUENCE OF 125-371 FROM N.A.
RX MEDLINE=92158051; PubMed=1741034;
RA Bankmann M., Prakash L., Prakash S.;
RT "Yeast RAD14 and human Xeroderma pigmentosum group A DNA-repair genes

```

```

encode homologous proteins."
RL Nature 355:555-558(1992).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=S288c / AB972;
RX MEDLINE=97311269; PubMed=9169872;
RA Bowman R., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA Jags K., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
RA Rice P., Skelton J., Walsh S., Whitehead S., Barrell B.G.;
RA "The nucleotide sequence of Saccharomyces cerevisiae chromosome
RT XIII."
RT Nature 387:90-93(1997).
RL Nature 387:90-93(1997).
RN [4]
RP CHARACTERIZATION.
RX Guzman S.N., Sung P., Prakash L., Prakash S.;
RA "Yeast DNA-repair gene RAD14 encodes a zinc metalloprotein with
RT affinity for ultraviolet-damaged DNA."
RL Proc. Natl. Acad. Sci. U.S.A. 90:5433-5437(1993).
CC -!- FUNCTION: INVOLVED IN NUCLEOTIDE EXCISION REPAIR. BINDS
CC SPECIFICALLY TO DAMAGED DNA. REQUIRED FOR THE INCISION STEP.
CC -----
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Belongs to the XPA family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X64064; CAA45420.1;
DR EMBL; Z48755; CAA88642.1;
DR F03722; S59442.
DR G0004814; RAD14.
DR SGD; S0004814; RAD14.
DR InterPro; IPR000465; XPA protein.
DR Pfam; PF05181; XPA_C; 1.
DR TIGRFAMs; TIGR00598; rad14; 1.
DR PROSITE; PS00752; XPA_1; 1.
DR PROSITE; PS00753; XPA_2; 1.
DR DNA repair; DNA-binding; Zinc-finger; Nuclear protein.
FT ZN FING 191 216
FT MUTAGEN 207 207
FT MUTAGEN 216 216
FT MUTAGEN 216 216
SQ SEQUENCE 371 AA; 43038 MW; A90FA9CAA6F3EA52 CRC64;

Query Match 56.0%; Score 42; DB 1; Length 371;
Best Local Similarity 63.6%; Pred. No. 6.8;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 4 KXDELRRKKWYM 14
Db 166 KXDELRRKKWYM 176

RESULT 4
REV HV2ST STANDARD; PRT; 107 AA.
ID REV HV2ST
AC P20870;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE REV protein (Anti-repression transactivator protein) (ART/TRS).
GN REV.
OS Human immunodeficiency virus type 2 (isolate ST) (HIV-2).
OC Viruses; Retroviridae; Retrovirus.
OX NCBI_TaxID=11721;
RN [1]

```


DR SMART; SM00360; RRM; 1.
 DR SMART; SM00356; ZNF_C3HL; 2.
 DR PROSITE; PS0102; RRM; 1.
 DR PROSITE; PS00030; RRM_RNP_1; FALSE NEG.
 KW Nuclear protein; RNA-binding; Ribonucleoprotein; Zinc-finger; Repeat.
 FT DOMAIN 202 308 RNA-BINDING (RRM).
 FT DOMAIN 50 53 POLY-GLU.
 FT DOMAIN 93 96 POLY-LYS.
 FT DOMAIN 122 127 POLY-GLU.
 FT DOMAIN 380 386 POLY-ARG.
 SQ SEQUENCE 462 AA; 55357 MW; 0E980C029ACBA196 CRC64;
 Query Match 50.7%; Score 38; DB 1; Length 462;
 Best Local Similarity 58.3%; Pred. No. 42;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 1 ARAKXDELPRKM 12
 DB 100 ARKKELEERKL 111
 RESULT 7
 T2F1_FLAOK STANDARD; PRT; 583 AA.
 AC P14870;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Type IIS restriction enzyme FokI (EC 3.1.21.4) (Endonuclease FokI)
 DE (A.FokI).
 GN FOKIR OR RFOKI.
 OS Flavobacterium okeanokoites.
 OC Bacteria; Firmicutes; Bacillales; Planococcaceae; Planomicrobium.
 OX NCBI_TaxID=244;
 RN [1]
 RP SEQUENCE FROM M.A.
 RC STRAIN=ATCC 33414 / IFO 12536;
 RX MEDLINE=89174630; PubMed=2784436;
 RA Kita K., Kotani H., Sugisaki H., Takanami M.;
 RT "The foki restriction-modification system. I. Organization and
 RT nucleotide sequences of the restriction and modification genes.";
 RL J. Biol. Chem. 264:5751-5756(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90060769; PubMed=2684765;
 RA Looney M.C., Moran L.S., Jack W.E., Feehery G.R., Benner J.S.,
 RA Slatko B.E., Wilson G.G.;
 RT "Nucleotide sequence of the FokI restriction-modification system:
 RT separate strand-specificity domains in the methyltransferase.";
 RL Gene 80:193-208(1989).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
 RC STRAIN=ATCC 33414 / IFO 12536;
 RX MEDLINE=97357159; PubMed=9214510;
 RA Wah D.A., Hirsch J.A., Dörner L.F., Schildkraut I., Aggarwal A.K.;
 RT "Structure of the multimodular endonuclease FokI bound to DNA.";
 RL Nature 388:97-100(1997).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
 RC STRAIN=ATCC 33414 / IFO 12536;
 RX MEDLINE=98393684; PubMed=9724743;
 RA Wah D.A., Bitinaite J., Schildkraut I., Aggarwal A.K.;
 RT "Structure of FokI has implications for DNA cleavage.";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:10564-10569(1998).
 RN [5]
 RP SUBUNIT.
 RX MEDLINE=98393685; PubMed=9724744;
 RA Bitinaite J., Wah D.A., Aggarwal A.K., Schildkraut I.;
 RT "FokI dimerization is required for DNA cleavage.";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:10570-10575(1998).
 CC -1- FUNCTION: RECOGNIZES THE DOUBLE-STRANDED SEQUENCES GGATG AND CATCC
 CC AND CLEAVES RESPECTIVELY 14 BASES AFTER G-1 AND 13 BASES BEFORE C-
 CC 1.

CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage of DNA to give
 CC specific double-stranded fragments with terminal 5'-phosphates.
 CC -1- COFACTOR: Magnesium.
 CC -1- SUBUNIT: Monomer. Homodimer when bound to DNA.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: J04623; AAA24927.1; .
 CC EMBL: M28828; AAA24934.1; ALT_INIT.
 CC PDB; 1FOK; 03-DEC-97.
 CC PDB; 2FOK; 17-JUN-98.
 CC REBASE; 1056; FOKI.
 CC InterPro; IPR004233; FOKI_C.
 CC InterPro; IPR004234; FOKI_N.
 CC Pfam; PF02980; FOKI_C; 1.
 CC Pfam; PF02981; FOKI_N; 1.
 KW Restriction system; Hydrolase; Nuclease; Endonuclease; Magnesium;
 KW 3D-structure.
 FT HELIX 21 28
 FT TURN 29 30
 FT TURN 33 34
 FT HELIX 36 43
 FT TURN 44 44
 FT HELIX 45 49
 FT HELIX 53 64
 FT STRAND 70 71
 FT HELIX 72 76
 FT HELIX 91 95
 FT HELIX 99 101
 FT HELIX 108 120
 FT TURN 121 122
 FT STRAND 124 127
 FT TURN 128 131
 FT STRAND 132 135
 FT HELIX 137 144
 FT TURN 147 148
 FT HELIX 151 160
 FT TURN 161 162
 FT HELIX 164 173
 FT TURN 174 175
 FT STRAND 179 180
 FT HELIX 181 185
 FT TURN 186 187
 FT TURN 191 192
 FT HELIX 194 196
 FT TURN 201 209
 FT TURN 210 210
 FT HELIX 213 215
 FT HELIX 216 222
 FT HELIX 226 240
 FT TURN 241 242
 FT STRAND 244 246
 FT STRAND 249 252
 FT STRAND 255 255
 FT TURN 256 257
 FT STRAND 258 258
 FT TURN 259 260
 FT STRAND 262 265
 FT STRAND 268 271
 FT STRAND 273 282
 FT TURN 283 284
 FT STRAND 293 293
 FT HELIX 297 297
 FT TURN 305 306
 FT HELIX 307 320
 FT TURN 321 322
 FT STRAND 327 328

```

T HELIX 329 338
T TURN 339 340
T TURN 345 357
T TURN 358 359
T STRAND 362 365
T TURN 366 367
T STRAND 368 371
T STRAND 375 375
T HELIX 393 404
T HELIX 410 414
T HELIX 415 418
T TURN 419 420
T TURN 422 423
T HELIX 425 438
T TURN 439 440
T STRAND 443 446
T STRAND 455 458
T STRAND 467 473
T STRAND 479 479
T HELIX 483 498
T TURN 501 503
T TURN 505 506
T HELIX 508 511
T TURN 514 515
T STRAND 519 525
T STRAND 528 528
T HELIX 532 543
T STRAND 546 551
T HELIX 552 564
T TURN 565 565
T HELIX 569 573
T TURN 574 575
T STRAND 580 580
T SEQUENCE 583 AA; 66219 MW; F4AF13DEFA5CF09 CRC64;

Query Match 50.7%; Score 38; DB 1; Length 583;
Best Local Similarity 58.3%; Pred. No. 53;
Matches 7; Mismatches 2; Indels 0; Gaps 0;

iy 4 KDELRRRQYIM 15
ib 397 KKSBLRHKLTV 408

RESULT 8
NAPA_ECOLI STANDARD; PRT; 828 AA.
AC P33937; P78087;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Periplasmic nitrate reductase precursor (EC 1.7.99.4).
EN NAPA OR B2206.
DS Escherichia coli.
DC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
CC Enterobacteriaceae; Escherichia.
CX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / BHB2600;
RA Richterich P., Lakey N., Gryan G., Jaehn L., Mintz L., Robison K.,
RA Church G.M.;
RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP REVISIONS.
RA Robison K.;
RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RA Blattner F.R.; Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.P.,
Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
Mau B., Shao Y.;
*The complete genome sequence of Escherichia coli K-12.*;
Science 277:1453-1474(1997).
[4]
SEQUENCE OF 1-757 FROM N.A.
RC STRAIN-K12;
MEDLINE=97251358; PubMed=9097040;
RA Itoh T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
RA Kasai H., Kimura S., Kitakawa M., Kitagawa M., Makino K., Miki T.,
RA Mizobuchi K., Mori H., Mori T., Motomura K., Nakade S., Nakamura Y.,
RA Nishimoto H., Nishio Y., Oshima T., Saito N., Sampei G., Seki Y.,
RA Sivasubdam S., Tagami H., Takeda J., Takemoto K., Wada C.,
RA Yamamoto Y., Horiuchi T.;
* A 460-kb DNA sequence of the Escherichia coli K-12 genome
corresponding to the 40.1-50.0 min region on the linkage map.*;
DNA Res. 3:379-392(1996)
-|- FUNCTION: LARGE SUBUNIT OF THE PERIPLASMIC NITRATE REDUCTASE
(NAP). ONLY EXPRESSED AT HIGH LEVELS DURING AEROBIC GROWTH. NAPAB
COMPLEX RECEIVES ELECTRONS FROM THE MEMBRANE-ANCHORED TETRAHEME
NAPC PROTEIN, THUS ALLOWING ELECTRON FLOW BETWEEN MEMBRANE AND
PERIPLASM. ESSENTIAL FUNCTION FOR NITRATE ASSIMILATION AND MAY
HAVE A ROLE IN ANAEROBIC METABOLISM (BY SIMILARITY).
-|- CATALYTIC ACTIVITY: Nitrite + acceptor = nitrate + reduced
acceptor.
-|- COFACTOR: Molybdenum (molybdopterin); may bind a 4Fe-4S cluster
(by similarity).
-|- SUBCELLULAR LOCATION: Periplasmic (Potential).
-|- SIMILARITY: Belongs to the prokaryotic molybdopterin-containing
oxidoreductase family. Napa/napa/narb subfamily.
-|- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO FRAMESHIFTS
IN POSITIONS 19 AND 27.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).

EMBL; U00008; AAA16399.1; ALT FRAME.
EMBL; AB000309; AAC75266.1; -.
EMBL; D90850; BAA15989.1; -.
PIR; D64990; D64990.
HSP; P81186; 2NRP.
Ecogene; EG1067; napa.
InterPro; IPR009010; Asp_decarb_fold.
InterPro; IPR006657; Mol_dinuc_bind.
InterPro; IPR006963; Molybdop_Fe4S4.
InterPro; IPR006656; Molybdopterin.
InterPro; IPR006655; Prok_Mboxred.
InterPro; IPR006311; Tat.
Pfam; PF04879; Molybdop_Fe4S4; 1.
Pfam; PF00384; Molybdopterin; 1.
Pfam; PF01568; Molybdop_binding; 1.
TIGRFAMs; TIGR01409; TAT_signal_seq; 1.
PROSITE; PS00551; MOLYBDOPTERIN_PROK_1; 1.
PROSITE; PS00490; MOLYBDOPTERIN_PROK_2; FALSE_NEG.
PROSITE; PS00932; MOLYBDOPTERIN_PROK_3; FALSE_NEG.
KW Nitrate assimilation; Oxidoreductase; Electron transport; Molybdenum;
KW Periplasmic; Signal; Iron-sulfur; 4Fe-4S; Complete proteome.
FT SIGNAL 1 31
FT CHAIN 32 828
FT METAL 46 46
FT METAL 49 49
FT METAL 53 53
FT METAL 81 81
FT CONFLICT 98 98
FT SEQUENCE 828 AA; 93042 MW; CE2D7AB000FEAFCA CRC64;
Query Match 50.7%; Score 38; DB 1; Length 828;
Best Local Similarity 66.7%; Pred. No. 76;

```

```

Matches      8;  Conservative      1;  Mismatches      3;  Indels      0;  Gaps      0;

QY      3 AKKDELRXOMY 14
      ||| ||| | | |
DB      581 AKKPELRGKTL 592

RESULT 9
THH2_PICMA
ID THH2_PICMA STANDARD; PRT; 125 AA.
AC 065049;
DF 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Thioresoxin H-type (TRX-H).
GN SB09.
OS Picea mariana (Black spruce).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Pinaceae; Picea.
OX NCBI_TaxID=3335;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=98278823; PubMed=9611216;
RA Perry D.J., Bousquet J.;
RT "Sequence-tagged-site (STS) markers of arbitrary genes: development,
RL Genetics 149:1089-1098(1998).
CC -!- FUNCTION: Participates in various redox reactions through the
CC reversible oxidation of the active center dithiol to a disulfide.
CC The H form is known to activate a number of cytosolic enzymes (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: BELONGS TO THE THIOREDOXIN FAMILY. PLANT H-TYPE.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF051206; AAC32111.1; -.
CC FIR; T50866; T50866.
CC HSSP; P10599; IERV.
CC InterPro; IPR006662; Thioresox_dom2.
CC PRINTS; PR00421; THIOREDOXIN.
CC PROSITE; PS00194; THIOREDOXIN; FALSE_NEG.
KW Redox-active center; Electron transport; Multigene family.
FT DISULFID 38 41 REDOX-ACTIVE (BY SIMILARITY).
SQ SEQUENCE 125 AA; 13732 MW; 37E1A22E0B5E229F CRC64;

Query Match      49.3%; Score 37; DB 1; Length 125;
Best Local Similarity 70.0%; Pred. No. 16;
Matches      7;  Conservative      2;  Mismatches      1;  Indels      0;  Gaps      0;

QY      3 AKKDELRXOM 12
      ||| | | | |
DB      98 AKKXDLERXV 107

RESULT 10
THH2_ARATH
ID THH2_ARATH STANDARD; PRT; 133 AA.
AC Q38879; Q39240;
DF 16-OCT-2001 (Rel. 40, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Thioresoxin H-type 2 (TRX-H-2).
GN TEX2 OR AT5G3950 OR MYH19.14.
OS Arabidopsis thaliana (Mouse-ear cress).

```

```

OC Rukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eusoids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN (1)
RP SEQUENCE FROM N.A.
RX STRAIN=cv. Columbia; TISSUE=Callus;
RA Rivera-Madrid R., Mestres D., Marinho P., Jacquot J.P.,
RA Decottignies P., Miginiac-Maslow M., Meyer Y.;
RT "Evidence for five divergent thioresoxin h sequences in Arabidopsis
thaliana.";
RN Proc. Natl. Acad. Sci. U.S.A. 92:5620-5624(1995).
RP SEQUENCE FROM N.A.
RX STRAIN=cv. Landsberg erecta;
RA Sahravy M., Hecht V., Lopez Jaramillo J., Chueca A., Chartier Y.,
RA Meyer Y.;
RT "Intron position as an evolutionary marker of thioresoxins and
thioresoxin domains.";
RN J. Mol. Evol. 42:422-431(1996).
RP SEQUENCE FROM N.A.
RX STRAIN=cv. Columbia;
RA Sato S., Kaneko T., Kotani H., Nakamura Y., Asamizu E., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. IV
Sequence features of the regions of 1,456,315 bp covered by nineteen
physically assigned P1 and TAC clones.";
RN DNA Res. 5:41-54(1998).
CC -!- FUNCTION: Participates in various redox reactions through the
CC reversible oxidation of the active center dithiol to a disulfide.
CC The H form is known to activate a number of cytosolic enzymes (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: BELONGS TO THE THIOREDOXIN FAMILY. PLANT H-TYPE.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z35475; CAA84612.1; -.
CC EMBL; U35826; AAC49353.1; -.
CC EMBL; AB010077; BAB10219.1; -.
CC FIR; S58123; S58123.
CC HSSP; P10599; IERV.
CC InterPro; IPR006662; Thioresox.
CC InterPro; IPR006663; Thioresox_dom2.
CC Pfam; PF00085; Thioresox; 1.
CC PRINTS; PR00421; THIOREDOXIN.
CC PROSITE; PS00194; THIOREDOXIN; 1.
KW Redox-active center; Electron transport; Multigene family.
FT DISULFID 59 62 REDOX-ACTIVE (BY SIMILARITY).
FT CONFLICT 7 7 T -> I (IN REF. 2).
FT CONFLICT 22 22 S -> SS (IN REF. 2).
FT CONFLICT 127 127 K -> Q (IN REF. 2).
SQ SEQUENCE 133 AA; 14676 MW; BC034E2BCD4D3CA3 CRC64;

Query Match      49.3%; Score 37; DB 1; Length 133;
Best Local Similarity 70.0%; Pred. No. 17;
Matches      7;  Conservative      2;  Mismatches      1;  Indels      0;  Gaps      0;

QY      3 AKKDELRXOM 12
      ||| | | | |
DB      119 AKKDELRXV 128

```

```
RESULT 11
DP3E_BUCBP STANDARD; PRT; 245 AA.
AC Q89AN3;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE DNA polymerase III, epsilon chain (EC 2.7.7.7).
GN DNAO OR BBP230.
OS Buchnera aphidicola (subsp. Baizongia pistaciae).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=135942;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426901; PubMed=1252265;
RA Van Ham R.C.H.J., Kamerbeek J., Palacios C., Rausell C., Abascal P.,
RA Bastolla U., Fernandez J.M., Jimenez L., Postigo M., Silva F.J.,
RA Tamames J., Viguera E., Latorre A., Valencia A., Moran F., Moya A.,
RT "Reductive genome evolution in Buchnera aphidicola.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:581-585 (2003).
CC -!- FUNCTION: DNA POLYMERASE III IS A COMPLEX, MULTICHAIN ENZYME
CC RESPONSIBLE FOR MOST OF THE REPLICATIVE SYNTHESIS IN BACTERIA.
CC THE EPSILON SUBUNIT CONTAIN THE EDITING FUNCTION AND IS A
CC PROOFREADING 3'-5' EXONUCLEASE (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + [DNA] (N).
CC -!- SUBUNIT: DNA polymerase III contains a core (composed of alpha,
CC epsilon and theta chains) that associates with a tau subunit. This
CC core dimerizes to form the POLIII' complex. POLIII' associates
CC with the gamma complex (composed of gamma, delta, delta', psi and
CC chi chains) and with the beta chain to form the complete DNA
CC polymerase III complex. The final composition of the complex is:
CC {alpha,epsilon,theta}[2]-{tau[2]-{gamma,delta,delta',psi,chi}[2]-
CC beta[4]} (By similarity).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; A5014016; A5026958.1; --
CC InterPro; IPR006055; Exonuclease.
CC Pfam; PF00929; Exonuclease; 1.
CC XW Transferase; DNA-directed DNA polymerase; DNA replication; Hydrolase;
CC XW Exonuclease; Complete proteome.
CC SEQUENCE 245 AA; 28519 MW; 866B40B552C5027B CRC64;
CC
CC Query Match 49.3%; Score 37; DB 1; Length 245;
CC Best Local Similarity 53.0%; Pred. No. 32;
CC Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
CC
CC YQ 3 AKKDELRFRQMTM 15
CC DB 215 ATKDELQHQMKYL 227
CC
CC RESULT 12
YDR8_SCHPO STANDARD; PRT; 269 AA.
AC O13743;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein C1658.08 in chromosome I.
GN SPAC1658.08.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
```

```
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Scouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Church C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Ganties S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy I., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skellon J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Voicakert G., Aert R., Robben J., Grymmonprez B.,
RA Weidjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
RA Beger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallard C., Fallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Carrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880 (2002).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; Z98529; CAB11036.1; --
CC PIR; T37788; T37788.
CC DR GeneDB SPombe; SPAC1658.08; --
CC KW Hypothetical protein.
CC SEQUENCE 269 AA; 30500 MW; 294F7864A310D16A CRC64;
CC
CC Query Match 49.3%; Score 37; DB 1; Length 269;
CC Best Local Similarity 46.2%; Pred. No. 36;
CC Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
CC
CC QY 2 RAKQDELRRKMY 14
CC DB 40 KAKENLEKLY 52
CC
CC RESULT 13
FSL1_MOUSE STANDARD; PRT; 306 AA.
ID FSL1_MOUSE
AC Q62356; Q99J19;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Follistatin-related protein 1 precursor (Follistatin-like 1) (TGF-
DE beta-inducible protein TSC-36).
GN FSTL1 OR FSTL OR FRP OR TSC36.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94039028; PubMed=7901004;
RA Shibamura M., Mashimo J., Mita A., Kuroki T., Nose K.;
```

RT "Cloning from a mouse osteoblastic cell line of a set of
RT transforming-growth-factor-beta 1-related genes, one of which
RL seems to encode a follistatin-regulated polypeptide."
RL Bur. J. Biochem. 217:13-19(1993).
RN [2]

RP SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wegner L., Shemmen C.M., Schuler G.D.,
RA Altschul S.F., Zebrow B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gamaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Munz D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra W.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

CC -1- FUNCTION: May modulate the action of some growth factors on cell
CC proliferation and differentiation. Binds heparin (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted (Potential).
CC -1- SIMILARITY: Contains 1 follistatin-like domain.
CC -1- SIMILARITY: Contains 1 Kzall-like domain.
CC -1- SIMILARITY: Contains 1 WFCC domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
CC EMBL; M91380; AAC37633.1; -;
CC EMBL; BC006185; -; NOT_ANNOTATED_CDS.
CC PIR; S38251; S38251.
CC HSP; P01001; ZBUS.
CC MGI; 102793; Fstl.
CC InterPro; IPR002048; EF-hand.
CC InterPro; IPR003645; FOLN.
CC InterPro; IPR002350; Kzall.
CC InterPro; IPR001007; WFCC.
CC Pfam; PF00036; efband; 2.
CC Pfam; PF00050; Kzall; 1.
CC SMART; SM00274; FOLN; 1.
CC SMART; SM00280; KZALL; 1.
CC PROSITE; PS00282; KZALL; FALSE NEG.
CC PROSITE; PS01208; WFCC 1; FALSE NEG.
CC Glycoprotein; signal; Heparin-binding.

CC SIGNAL 1 18 BY SIMILARITY.
CC CHAIN 19 306 FOLLISTATIN-RELATED PROTEIN 1.
CC DOMAIN 28 51 FOLLISTATIN.
CC DOMAIN 52 96 KZALL-LIKE.
CC DOMAIN 231 285 WFCC.
CC DISULFID 52 82 BY SIMILARITY.
CC DISULFID 56 75 BY SIMILARITY.
CC DISULFID 64 96 BY SIMILARITY.
CC CARBOHYD 142 142 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 173 173 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 178 178 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CONFLICT 235 235 V -> D (IN REF. 2).
CC SEQUENCE 306 AA; 34538 MW; 4631070AEB1E8FC4 CRC64;

Query Match 49.3%; Score 37; DB 1; Length 306;

Best Local Similarity 50.0%; Pred. No. 41;
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
QY 2 RAKKDELRRKRM 13
DB 113 QANRDELRRRLI 124
:|||||::
:|||||::

RESULT 14
HSLU THEMA STANDARD; PRT; 463 AA.
AC QWY2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE ATP-dependent hsl protease ATP-binding subunit hslu.
GN HSLU OR TM0522;
OS Thermotoga maritima.
OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]

RP SEQUENCE FROM N.A.
RX STRAIN=MSB / DSM 3109 / ATCC 43589;
RX MEDLINE=9287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Haft D.H., Hickley E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Lither K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter C.M., Fraser C.M.;
RT "Evidence for lateral gene transfer: between Archaea and Bacteria from
RT genome sequence of Thermotoga maritima."
RL Nature 391:323-329(1999).

CC -1- FUNCTION: Chaperone subunit of a proteasome-like degradation
CC complex (By similarity).
CC -1- SUBUNIT: A double ring-shaped homohexamer of hslv is capped on
CC each side by a ring-shaped hslu homohexamer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the clpX chaperone family. Hslu subfamily.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
CC EMBL; AE001728; AAD35607.1; -;
CC PIR; H72365; H72365.
CC HSP; P32168; 1E94.
CC TIGR; TM0522; -;
CC HAMAP; MF_00249; -; 1.
CC InterPro; IPR003593; AAA ATPase.
CC InterPro; IPR003959; AAA ATPase_centre.
CC InterPro; IPR004491; Hsp_HsluV.
CC Pfam; PF00004; AAA; 2.
CC SMART; SM00382; AAA; 1.
CC TIGREMS; TIGR00390; hslu; 1.
CC Chaperone; ATP-binding; Complete proteome.

CC NP_BIND 60 67 ATP (POTENTIAL).
CC FT SEQUENCE 463 AA; 53052 MW; F871CD909FCBA5CA CRC64;

Query Match 49.3%; Score 37; DB 1; Length 463;
Best Local Similarity 54.5%; Pred. No. 62;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 RAKKDELRRKRM 12
DB 173 RAKREEMREKL 183
|||||:
|||||:

RESULT 15

```

JFPA TREPA
ID_CFFA TREPA STANDARD; PRT; 677 AA.
IC_Q56336;
IT_15-DEC-1998 (Rel. 37, Created)
JT_30-MAY-2000 (Rel. 39, Last sequence update)
JE_16-OCT-2001 (Rel. 40, Last annotation update)
DE_Cytoplasmic filament protein A.
FN_CFFA OR TPN83 OR TP0748.
NS_Treponema pallidum.
XC_Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
XX_NCBITaxID=160;
XN_[1]
P_SEQUENCE FROM N.A., AND SEQUENCE OF 1-21 AND 146-161.
IC_STRAIN=Nichols;
X_MEDLINE=96236033; PubMed=8655496;
A_You Y., Elmore S., Colton L.L., Mackenzie C., Stoops J.K.,
T_Weinstock G.M., Norris S.J.;
T_"Characterization of the cytoplasmic filament protein gene (cfpa) of
T_Treponema pallidum subsp. pallidum.";
L_J. Bacteriol. 178:3177-3187(1996).
XN_[2]
P_SEQUENCE FROM N.A.
IC_STRAIN=Nichols;
X_MEDLINE=98332770; PubMed=9655876;
A_Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
A_Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
A_Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
A_Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
A_McDonald L., Artiach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
A_Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
A_Venter J.C.;
T_"Complete genome sequence of Treponema pallidum, the syphilis
T_spirochete.";
L_Science 281:375-388(1998).
C_-1- FUNCTION: COMPONENT OF THE CYTOPLASMIC FILAMENTS THAT RUN THE
C_LENGTH OF THE ORGANISM JUST UNDERNEATH THE CYTOPLASMIC MEMBRANE.
C_-1- SUBCELLULAR LOCATION: AN ARRAY OF 4 TO 6 FILAMENTS LIE IN CLOSE
C_APPPOSITION TO THE INNER MEMBRANE AND ARE ALWAYS LOCALIZED DIRECTLY
C_UNDERNEATH THE CORRESPONDING GROUP OF PERIPLASMIC FLAGELLA.
C-----
C_This SWISS-PROT entry is copyright. It is produced through a collaboration
C_between the Swiss Institute of Bioinformatics and the EMBL outstation -
C_the European Bioinformatics Institute. There are no restrictions on its
C_use by non-profit institutions as long as its content is in no way
C_modified and this statement is not removed. Usage by and for commercial
C_entities requires a license agreement (See http://www.isb-sib.ch/announce/
C_or send an email to license@isb-sib.ch).
C-----
R_EMBL; U32683; AAB07757.1; -.
R_EMBL; AE001246; AAC65715.1; -.
R_PIR; A71287; A71287.
R_TIGR; TP0748; -.
M_Structural protein; Antigen; Coiled coil; Complete proteome.
T_INIT MET 0
T_DOMAIN 23 48 COILED COIL (POTENTIAL).
T_DOMAIN 289 310 COILED COIL (POTENTIAL).
T_DOMAIN 396 416 COILED COIL (POTENTIAL).
T_DOMAIN 521 539 COILED COIL (POTENTIAL).
T_SEQUENCE 677 AA; 78408 MW; 7F77A144214B7948 CRC64;
Q_QUERY Match 49.3%; Score 37; DB 1; Length 677;
Q_Best local Similarity 50.0%; Pred. No. 92;
Q_Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
Y_4 KDELRRKMMYM 15
b_495 EKDKMRKILM 506

```

earch completed: May 21, 2004, 12:47:59
 ob time : 14.6957 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

WM protein - protein search, using sw model

Run on: May 21, 2004, 12:43:34 ; Search time 47.6087 Seconds
(without alignments)
99.410 Million cell updates/sec

Title: US-09-980-058-2

Perfect score: 75

Sequence: 1 ARAKDELRRKMMYM 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

result No.	Score	Query Match	Length	DB ID	Description
1	75	100.0	100	12 Q9W9C4	Q9W9C4 human cytom
2	75	100.0	100	12 Q9YRL5	Q9YRL5 human cytom
3	75	100.0	104	12 Q9WKS1	Q9WKS1 human cytom
4	75	100.0	105	12 Q9WKS7	Q9WKS7 human cytom
5	75	100.0	107	12 Q9YRM0	Q9YRM0 human cytom
6	75	100.0	107	12 Q9WKS2	Q9WKS2 human cytom
7	75	100.0	108	12 Q9YRK7	Q9YRK7 human cytom
8	75	100.0	111	12 Q9W8Q5	Q9W8Q5 human cytom
9	72	96.0	100	12 Q9YRL4	Q9YRL4 human cytom
10	72	96.0	109	12 Q9YEM2	Q9YEM2 human cytom
11	72	96.0	110	12 Q9YRM1	Q9YRM1 human cytom
12	72	96.0	112	12 Q9YRL2	Q9YRL2 human cytom
13	72	96.0	125	12 Q9YRK3	Q9YRK3 human cytom
14	71	94.7	99	12 Q9YRK5	Q9YRK5 human cytom
15	71	94.7	100	12 Q9YRL1	Q9YRL1 human cytom
16	71	94.7	101	12 Q9WKR9	Q9WKR9 human cytom

17	71	94.7	105	12 Q9YRL7	Q9YRL7 human cytom
18	71	94.7	105	12 Q9YRL9	Q9YRL9 human cytom
19	71	94.7	110	12 Q9YRK9	Q9YRK9 human cytom
20	71	94.7	111	12 Q9YRK6	Q9YRK6 human cytom
21	68	90.7	100	12 Q9W8G6	Q9W8G6 human cytom
22	68	90.7	103	12 Q9WKS4	Q9WKS4 human cytom
23	68	90.7	105	12 Q9WKS6	Q9WKS6 human cytom
24	68	90.7	105	12 Q9YRK2	Q9YRK2 human cytom
25	68	90.7	105	12 Q9YRK1	Q9YRK1 human cytom
26	68	90.7	108	12 Q9WKS5	Q9WKS5 human cytom
27	68	90.7	108	12 Q9WKS0	Q9WKS0 human cytom
28	68	90.7	110	12 Q9YRL3	Q9YRL3 human cytom
29	68	90.7	110	12 Q9W9A8	Q9W9A8 human cytom
30	68	90.7	111	12 Q9YJL1	Q9YJL1 human cytom
31	67	89.3	100	12 Q9YRL0	Q9YRL0 human cytom
32	67	89.3	103	12 Q9YRK8	Q9YRK8 human cytom
33	66	88.0	100	12 Q9WKS3	Q9WKS3 human cytom
34	66	88.0	106	12 Q9YRK4	Q9YRK4 human cytom
35	66	88.0	106	12 Q9YRK0	Q9YRK0 human cytom
36	66	88.0	111	12 Q9YRL6	Q9YRL6 human cytom
37	66	88.0	111	12 Q9YRL8	Q9YRL8 human cytom
38	65	86.7	101	12 Q9YRK9	Q9YRK9 human cytom
39	52	69.3	505	12 Q8ORY6	Q8ORY6 chimpanzee
40	46	61.3	152	2 P72493	P72493 streptococo
41	46	61.3	528	16 Q98CJ3	Q98CJ3 rhizobium l
42	44	58.7	349	16 Q8RGW6	Q8RGW6 fusbacteri
43	43	57.3	68	16 Q8KC22	Q8KC22 chlorobium
44	43	57.3	521	16 Q81L97	Q81L97 bacillus an
45	42	56.0	141	16 Q9PJ24	Q9PJ24 campylobact

ALIGNMENTS

RESULT 1

Q9W9C4 PRELIMINARY; PRT; 100 AA.
ID Q9W9C4;
AC Q9W9C4;
DE 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, last annotation update)
DE Major immediate-early protein (Fragment).
OS Human cytomegalovirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10359;
RN 111
RP SEQUENCE FROM N.A.
RC STRAIN=1511/93, 574/85, 1614/84, and 79/90;
RA Zwegberg Wargart B., Brytting M., Linde A., Wahren B., Grillner L.;
RT "Sequence variation within three important cytomegalovirus gene regions in isolates from four different patient populations.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF099591; AAD04513.1; -;
DR EMBL; AF099553; AAD04475.1; -;
DR EMBL; AF099560; AAD04482.1; -;
DR EMBL; AF099576; AAD04498.1; -;
FT NON TER 1 1
FT NON TER 100 100
SQ SEQUENCE 100 AA; 11420 MW; 77AE0F88C1878C00 CRC64;

Query Match 100.0%; Score 75; DB 12; Length 100;
Best Local Similarity 100.0%; Pred. No. 5.5e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ARAKDELRRKMMYM 15

Db 45 ARAKDELRRKMMYM 59

RESULT 2

Q9YRL5 PRELIMINARY; PRT; 100 AA.
ID Q9YRL5

AC Q9YRL5;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Major immediate-early protein (Fragment).
 OS Human cytomegalovirus.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Cytomegalovirus.
 OX NCBI_TaxID=10359;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=77/90;
 RX MEDLINE=99036778; PubMed=9817892;
 RA Zwegberg Wirgart B., Brytting M., Linde A., Wahren B., Grillner L.;
 RT "Sequence variation within three important cytomegalovirus gene
 regions in isolates from four different patient populations.";
 RL J. Clin. Microbiol. 36:3662-3669(1998).
 DR EMBL; AF099587; AAD04509.1; -.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR002198; ADH short.
 DR PROSITE; PS00061; ADH SHORT; 1.
 FT NON_TER 1
 FT NON_TER 100 100
 SQ SEQUENCE 100 AA; 11459 MW; 6417704E7E259D14 CRC64;

Query Match 100.0%; Score 75; DB 12; Length 100;
 Best Local Similarity 100.0%; Pred. No. 5.8e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ARAKDELRRKKMYM 15
 Db 45 ARAKDELRRKKMYM 59

RESULT 3

ID Q9WKS1 PRELIMINARY; PRT; 104 AA.
 AC Q9WKS1;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Major immediate-early protein (Fragment).
 OS Human cytomegalovirus.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Cytomegalovirus.
 OX NCBI_TaxID=10359;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=454/94;
 RX MEDLINE=99036778; PubMed=9817892;
 RA Zwegberg Wirgart B., Brytting M., Linde A., Wahren B., Grillner L.;
 RT "Sequence variation within three important cytomegalovirus gene
 regions in isolates from four different patient populations.";
 RL J. Clin. Microbiol. 36:3662-3669(1998).
 DR EMBL; AF099587; AAD04509.1; -.
 DR NON_TER 1
 FT NON_TER 104 104
 SQ SEQUENCE 104 AA; 11911 MW; D64A1E8B47AE0F88 CRC64;

Query Match 100.0%; Score 75; DB 12; Length 104;
 Best Local Similarity 100.0%; Pred. No. 5.8e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ARAKDELRRKKMYM 15
 Db 45 ARAKDELRRKKMYM 59

RESULT 4

ID Q9WKS7 PRELIMINARY; PRT; 105 AA.
 AC Q9WKS7;

DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Major immediate-early protein (Fragment).
 OS Human cytomegalovirus.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Cytomegalovirus.
 OX NCBI_TaxID=10359;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1606/84;
 RX MEDLINE=99036778; PubMed=9817892;
 RA Zwegberg Wirgart B., Brytting M., Linde A., Wahren B., Grillner L.;
 RT "Sequence variation within three important cytomegalovirus gene
 regions in isolates from four different patient populations.";
 RL J. Clin. Microbiol. 36:3662-3669(1998).
 DR EMBL; AF099558; AAD04480.1; -.
 DR NON_TER 1
 FT NON_TER 105 105
 SQ SEQUENCE 105 AA; 12024 MW; B5664A1E8B47AE0F CRC64;

Query Match 100.0%; Score 75; DB 12; Length 105;
 Best Local Similarity 100.0%; Pred. No. 5.8e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ARAKDELRRKKMYM 15
 Db 45 ARAKDELRRKKMYM 59

RESULT 5

ID Q9YRMO PRELIMINARY; PRT; 107 AA.
 AC Q9YRMO;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Major immediate-early protein (Fragment).
 OS Human cytomegalovirus.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Cytomegalovirus.
 OX NCBI_TaxID=10359;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1577/84;
 RX MEDLINE=99036778; PubMed=9817892;
 RA Zwegberg Wirgart B., Brytting M., Linde A., Wahren B., Grillner L.;
 RT "Sequence variation within three important cytomegalovirus gene
 regions in isolates from four different patient populations.";
 RL J. Clin. Microbiol. 36:3662-3669(1998).
 DR EMBL; AF099557; AAD04479.1; -.
 DR NON_TER 1
 FT NON_TER 107 107
 SQ SEQUENCE 107 AA; 12315 MW; DF1085664A1E8B5F CRC64;

Query Match 100.0%; Score 75; DB 12; Length 107;
 Best Local Similarity 100.0%; Pred. No. 5.9e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ARAKDELRRKKMYM 15
 Db 45 ARAKDELRRKKMYM 59

RESULT 6

ID Q9WKS2 PRELIMINARY; PRT; 107 AA.
 AC Q9WKS2;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Major immediate-early protein (fragment).
 OS Human cytomegalovirus.

XC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 XC Betaherpesvirinae; Cytomegalovirus.
 XX NCBI_TaxID=10359;
 XN [1]
 XP SEQUENCE FROM N.A.
 XC STRAIN=829/91;
 XC MEDLINE=99036778; PubMed=9817892;
 XA Zwegberg Wirgart B., Brytting M., Linde A., Wahren B., Grillner L.;
 XT "Sequence variation within three important cytomegalovirus gene
 regions in isolates from four different patient populations.";
 RL J. Clin. Microbiol. 36:3662-3669(1998).
 DR EMBL; AF099578; AAD04500.1; -.
 FT NON_TER 1
 FT NON_TER 107
 SQ SEQUENCE 107 AA; 12299 MW; 85108564AJE8B47 CRC64;

Query Match 100.0%; Score 75; DB 12; Length 107;
 Best Local Similarity 100.0%; Pred. No. 5.9e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 ARAKKDELRKRM 15
 Y 45 ARAKKDELRKRM 59
 Y |||||

RESULT 7

ID Q9YRK7 PRELIMINARY; PRT; 108 AA.
 YC Q9YRK7
 YC 01-MAY-1999 (TReMBLrel. 10, Created)
 YC 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
 YC 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 YC Major immediate-early protein (Fragment).
 YC Human cytomegalovirus.
 XC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 XC Betaherpesvirinae; Cytomegalovirus.
 XX NCBI_TaxID=10359;
 XN [1]
 XP SEQUENCE FROM N.A.
 XC STRAIN=61/92;
 XC MEDLINE=99036778; PubMed=9817892;
 XA Zwegberg Wirgart B., Brytting M., Linde A., Wahren B., Grillner L.;
 XT "Sequence variation within three important cytomegalovirus gene
 regions in isolates from four different patient populations.";
 RL J. Clin. Microbiol. 36:3662-3669(1998).
 DR EMBL; AF099582; AAD04504.1; -.
 FT NON_TER 1
 FT NON_TER 108
 SQ SEQUENCE 108 AA; 12498 MW; 3F37186E997FDEC9 CRC64;

Query Match 100.0%; Score 75; DB 12; Length 108;
 Best Local Similarity 100.0%; Pred. No. 6e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 ARAKKDELRKRM 15
 Y 45 ARAKKDELRKRM 59
 Y |||||

RESULT 8

ID Q9W8Q5 PRELIMINARY; PRT; 111 AA.
 YC Q9W8Q5
 YC 01-NOV-1999 (TReMBLrel. 12, Created)
 YC 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
 YC 01-OCT-2001 (TReMBLrel. 18, Last annotation update)
 YC Major immediate-early protein (Fragment).
 YC Human cytomegalovirus.
 XC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 XC Betaherpesvirinae; Cytomegalovirus.
 XX NCBI_TaxID=10359;
 XN [1]
 XP SEQUENCE FROM N.A.

RC STRAIN=241/94, 1366/90, and 185/85;
 RA Zwegberg Wirgart B., Brytting M., Linde A., Wahren B., Grillner L.;
 RT "Sequence variation within three important cytomegalovirus gene regions
 in isolates from four different patient populations.";
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF099593; AAD04515.1; -.
 DR EMBL; AF099584; AAD04486.1; -.
 DR EMBL; AF099586; AAD04488.1; -.
 FT NON_TER 1
 FT NON_TER 111
 SQ SEQUENCE 111 AA; 12770 MW; 09E95018A5108566 CRC64;

Query Match 100.0%; Score 75; DB 12; Length 111;
 Best Local Similarity 100.0%; Pred. No. 6.1e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARAKKDELRKRM 15
 Db 45 ARAKKDELRKRM 59
 QY |||||

RESULT 9

ID Q9YRL4 PRELIMINARY; PRT; 100 AA.
 AC Q9YRL4
 DT 01-MAY-1999 (TReMBLrel. 10, Created)
 DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE Major immediate-early protein (Fragment).
 OS Human cytomegalovirus.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Cytomegalovirus.
 OX NCBI_TaxID=10359;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=979/91;
 RX MEDLINE=99036778; PubMed=9817892;
 RA Zwegberg Wirgart B., Brytting M., Linde A., Wahren B., Grillner L.;
 RT "Sequence variation within three important cytomegalovirus gene
 regions in isolates from four different patient populations.";
 RL J. Clin. Microbiol. 36:3662-3669(1998).
 DR EMBL; AF099573; AAD04495.1; -.
 FT NON_TER 1
 FT NON_TER 100
 SQ SEQUENCE 100 AA; 11431 MW; 4753A4509F9DA86F CRC64;

Query Match 96.0%; Score 72; DB 12; Length 100;
 Best Local Similarity 93.3%; Pred. No. 0.00018;
 Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARAKKDELRKRM 15
 Db 45 ARAKKDELRKRM 59
 QY |||||

RESULT 10

ID Q9YRM2 PRELIMINARY; PRT; 109 AA.
 AC Q9YRM2
 DT 01-MAY-1999 (TReMBLrel. 10, Created)
 DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE Major immediate-early protein (Fragment).
 OS Human cytomegalovirus.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Cytomegalovirus.
 OX NCBI_TaxID=10359;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=556/85;
 RX MEDLINE=99036778; PubMed=9817892;
 RA Zwegberg Wirgart B., Brytting M., Linde A., Wahren B., Grillner L.;
 RT "Sequence variation within three important cytomegalovirus gene

RT regions in isolates from four different patient populations.;"

RL J. Clin. Microbiol. 36:3662-3669(1998).

DR EMBL; AF099552; AAD04474.1; --

FT NON_TER 1

FT NON_TER 109

SQ SEQUENCE 109 AA; 12478 MW; 8E51F017D9D65FED CRC64;

Query Match 96.0%; Score 72; DB 12; Length 109;

Best Local Similarity 93.3%; Pred. No. 0.00019;

Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARAKKDELRRKKMYM 15

|||||:|||||

Db 45 ARAKKDELRRKKMYM 59

RESULT 11

ID Q9YRM1

AC Q9YRM1 PRELIMINARY; PRT; 110 AA.

DT 01-MAY-1999 (TReMBLrel. 10, Created)

DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)

DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)

DE Major immediate-early protein (Fragment).

OS Human cytomegalovirus.

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;

OC Betaherpesvirinae; Cytomegalovirus.

OX NCBI_TaxID=10359;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=605/85;

RX MEDLINE=99036778; PubMed=9817892;

RA Zweyberg Wirgart B., Brytting M., Linde A., Wahren B., Grillner L.;

RT "Sequence variation within three important cytomegalovirus gene

regions in isolates from four different patient populations.;"

RL J. Clin. Microbiol. 36:3662-3669(1998).

DR EMBL; AF099552; AAD04477.1; --

FT NON_TER 1

FT NON_TER 110

SQ SEQUENCE 110 AA; 12627 MW; E2BE51F01939D65F CRC64;

Query Match 96.0%; Score 72; DB 12; Length 110;

Best Local Similarity 93.3%; Pred. No. 0.00019;

Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARAKKDELRRKKMYM 15

|||||:|||||

Db 45 ARAKKDELRRKKMYM 59

RESULT 12

ID Q9YRL2

AC Q9YRL2 PRELIMINARY; PRT; 112 AA.

DT 01-MAY-1999 (TReMBLrel. 10, Created)

DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)

DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)

DE Major immediate-early protein (Fragment).

OS Human cytomegalovirus.

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;

OC Betaherpesvirinae; Cytomegalovirus.

OX NCBI_TaxID=10359;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=53/92;

RX MEDLINE=99036778; PubMed=9817892;

RA Zweyberg Wirgart B., Brytting M., Linde A., Wahren B., Grillner L.;

RT "Sequence variation within three important cytomegalovirus gene

regions in isolates from four different patient populations.;"

RL J. Clin. Microbiol. 36:3662-3669(1998).

DR EMBL; AF099575; AAD04497.1; --

FT NON_TER 1

FT NON_TER 112

SQ SEQUENCE 112 AA; 12815 MW; D5AE79C581F01939 CRC64;

Query Match 96.0%; Score 72; DB 12; Length 112;

Best Local Similarity 93.3%; Pred. No. 0.0002;

Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARAKKDELRRKKMYM 15

|||||:|||||

Db 45 ARAKKDELRRKKMYM 59

RESULT 13

ID Q9YRK3

AC Q9YRK3 PRELIMINARY; PRT; 125 AA.

DT 01-MAY-1999 (TReMBLrel. 10, Created)

DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)

DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)

DE Major immediate-early protein (Fragment).

OS Human cytomegalovirus.

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;

OC Betaherpesvirinae; Cytomegalovirus.

OX NCBI_TaxID=10359;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=249/94;

RX MEDLINE=99036778; PubMed=9817892;

RA Zweyberg Wirgart B., Brytting M., Linde A., Wahren B., Grillner L.;

RT "Sequence variation within three important cytomegalovirus gene

regions in isolates from four different patient populations.;"

RL J. Clin. Microbiol. 36:3662-3669(1998).

DR EMBL; AF099589; AAD04511.1; --

FT NON_TER 1

FT NON_TER 125

SQ SEQUENCE 125 AA; 14414 MW; 01C19770CFCE261 CRC64;

Query Match 96.0%; Score 72; DB 12; Length 125;

Best Local Similarity 93.3%; Pred. No. 0.00022;

Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARAKKDELRRKKMYM 15

|||||:|||||

Db 62 ARAKKDELRRKKMYM 76

RESULT 14

ID Q9YRK5

AC Q9YRK5 PRELIMINARY; PRT; 99 AA.

DT 01-MAY-1999 (TReMBLrel. 10, Created)

DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)

DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)

DE Major immediate-early protein (Fragment).

OS Human cytomegalovirus.

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;

OC Betaherpesvirinae; Cytomegalovirus.

OX NCBI_TaxID=10359;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=1343/93;

RX MEDLINE=99036778; PubMed=9817892;

RA Zweyberg Wirgart B., Brytting M., Linde A., Wahren B., Grillner L.;

RT "Sequence variation within three important cytomegalovirus gene

regions in isolates from four different patient populations.;"

RL J. Clin. Microbiol. 36:3662-3669(1998).

DR EMBL; AF099584; AAD04506.1; --

DR GO; GO:0016491; P:oxidoreductase activity; IEA.

DR GO; GO:0008152; P:metabolism; IEA.

DR InterPro; IPR002198; ADH short.

DR PROSITE; PS00061; ADH SHORT; 1.

FT NON_TER 1

FT NON_TER 99

SQ SEQUENCE 99 AA; 11420 MW; B425FEF4D53F28B4 CRC64;

Query Match 94.7%; Score 71; DB 12; Length 99;
 Best Local Similarity 100.0%; Pred.No. 0.00026;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RAKKDELRRKKMYM 15
 Db 45 RAKKDELRRKKMYM 58

RESULT 15

Q9YRL1 PRELIMINARY; PRT; 100 AA.
 AC Q9YRL1
 DT 01-MAY-1999 (TREMELrel. 10, Created)
 DT 01-MAY-1999 (TREMELrel. 10, Last sequence update)
 DE 01-DEC-2001 (TREMELrel. 19, Last annotation update)
 OS Major immediate-early protein (Fragment).
 OS Human cytomegalovirus.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Cytomegalovirus.
 OX NCBI_TaxID=10359;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=589/51;
 RX MEDLINE=99036778; PubMed=9817892;
 RA Zwegberg Wirgart B., Brytting M., Linde A., Wahren B., Grillner L.;
 RT "Sequence variation within three important cytomegalovirus gene
 regions in isolates from four different patient populations.";
 RL J. Clin. Microbiol. 36:3662-3669(1998).
 DR EMBL; AF099577; AAD0499.1; -.
 FT NON_TER 1
 FT NON_TER 100
 SQ SEQUENCE 100 AA; 11402 MW; 77AB4A9C95878C00 CRC64;

Query Match 94.7%; Score 71; DB 12; Length 100;
 Best Local Similarity 93.3%; Pred.No. 0.00026;
 Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARAKKDELRRKKMYM 15
 Db 45 ARAKKDELRRKKMYI 59

Search completed: May 21, 2004, 12:49:26
 Job time : 49.6087 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

CM protein - protein search, using sw model

Run on: May 21, 2004, 12:34:03 ; Search time 35.1304 Seconds
(without alignments)
64,343 Million cell updates/sec

Title: US-09-980-058-9_COPY_1_8

Perfect score: 41

Sequence: 1 ELRSXMY 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04.*

- 1: Geneseq1980s.*
- 2: Geneseq1990s.*
- 3: Geneseq2000s.*
- 4: Geneseq2001s.*
- 5: Geneseq2002s.*
- 6: Geneseq2003as.*
- 7: Geneseq2003bs.*
- 8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	9	4	Ab46271 HCMV IE-1
2	41	100.0	9	4	Ab46272 HCMV IE-1
3	41	100.0	9	4	Ab46142 HCMV type
4	41	100.0	9	4	Ab46141 HCMV type
5	41	100.0	9	6	Abj38054 Human cyt
6	41	100.0	10	4	Ab46265 HCMV IE-1
7	41	100.0	10	4	Ab46146 HCMV type
8	41	100.0	10	4	Ab46135 HCMV type
9	41	100.0	15	4	Ab465600 Human cyt
10	41	100.0	15	4	Ab46266 HCMV IE-1
11	41	100.0	15	4	Ab46264 HCMV IE-1
12	41	100.0	15	4	Ab46134 HCMV type
13	41	100.0	15	4	Ab46136 HCMV type
14	41	100.0	15	4	Ab46133 HCMV type
15	41	100.0	406	2	Ab46133 HCMV IE-e
16	41	100.0	406	2	Ab46133 HCMV IE-e
17	41	100.0	491	4	Ab46133 HCMV IE-e
18	38	92.7	9	6	Abj38056 Human cyt
19	38	92.7	241	6	Abj38259 Human cyt
20	38	92.7	241	6	Abj38258 Human cyt
21	38	92.7	241	6	Abj38252 Human cyt
22	38	92.7	241	6	Abj38257 Human cyt
23	38	92.7	241	6	Abj38260 Human cyt
24	38	92.7	241	6	Abj38256 Human cyt
25	38	92.7	241	6	Abj38254 Human cyt

26	38	92.7	241	6	Abj38251 Human cyt
27	38	92.7	241	6	Abj38255 Human cyt
28	38	92.7	250	6	Abj38253 Human cyt
29	33	80.5	109	4	Ab46141 HCMV type
30	32	78.0	9	6	Abj38057 Human cyt
31	32	78.0	15	4	Ab46146 HCMV type
32	32	78.0	20	6	Abj38062 Human cyt
33	31	75.6	222	3	Ab46141 HCMV type
34	31	75.6	430	6	Ab46141 HCMV type
35	31	75.6	803	7	Ab46141 HCMV type
36	31	75.6	1024	5	Ab46141 HCMV type
37	31	75.6	1286	5	Ab46141 HCMV type
38	31	75.6	1330	7	Ab46141 HCMV type
39	30	73.2	80	5	Ab46141 HCMV type
40	30	73.2	256	3	Ab46141 HCMV type
41	30	73.2	256	6	Ab46141 HCMV type
42	30	73.2	276	5	Ab46141 HCMV type
43	30	73.2	308	4	Ab46141 HCMV type
44	30	73.2	308	4	Ab46141 HCMV type
45	30	73.2	308	6	Ab46141 HCMV type

ALIGNMENTS

RESULT 1
AAB46271
ID AAB46271 standard; peptide; 9 AA.
XX AAB46271;
AC AAB46271;
XX
DT 04-APR-2001 (first entry)
XX
DE HCMV IE-1 antigenic peptide fragment SEQ ID NO 10.
XX
KW Antigen; vaccine; human; infection; Immediate-Early 1 Protein; IE-1;
KW Interferon-gamma; CD8+ cell; immunogenic.
XX
OS Human cytomegalovirus.
XX
PN DE19927039-A1.
XX
PD 07-DEC-2000.
XX
PF 04-JUN-1999; 99DE-01027039.
XX
PR 04-JUN-1999; 99DE-01027039.
PR 07-SEP-1999; 99DE-01043702.
XX
(KERN/) KERN F.
XX Kern F, Volk H, Reinke P, Faulhaber N, Surel I, Khatazias E;
XX MPI; 2001-061700/07.
XX
PT New peptides that stimulate cytokine production in T cells, useful for
PT vaccination against human cytomegalovirus infection and for detecting
PT immune responses to the virus.
XX
PS Claim 2; Page 8; 10pp; German.
XX
CC This invention describes novel peptides (including Immediate-Early 1
CC protein (IE-1) fragments) and their derivatives resulting from deletions,
CC insertions or substitutions which are capable of inducing the secretion
CC of interferon-gamma in CD8+ cells. The products of the invention have
CC immunogenic activity and are useful for the production of vaccines
CC against human cytomegaloviruses (HCMV) infections and for diagnostics for
CC the identification of HCMV infections
XX
SQ Sequence 9 AA;
Query Match 100.0%; Score 41; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ELRRKQMY 8
 OS |||||
 DB 1 ELRRKQMY 8

RESULT 2
 AAB46272
 ID AAB46272 standard; peptide; 9 AA.
 XX
 AC AAB46272;
 XX
 DT 04-APR-2001 (first entry)
 XX
 DE HCMV IE-1 antigenic peptide fragment SEQ ID NO 11.
 XX
 KW Antigen; vaccine; human; infection; Immediate-Early 1 Protein; IE-1;
 KW interferon-gamma; CD8+ cell; immunogenic.
 XX
 CS Human cytomegalovirus.
 XX
 PN DE19927039-A1.
 XX
 PD 07-DEC-2000.
 XX
 PF 04-JUN-1999; 99DE-01027039.
 XX
 PR 04-JUN-1999; 99DE-01027039.
 XX
 PR 07-SEP-1999; 99DE-01043702.
 XX
 PA (KERN/) KERN F.
 XX
 PI Kern F, Volk H, Reinke P, Faulhaber N, Surel I, Khatanzas E;
 XX
 DR WPI; 2001-061700/07.
 XX
 PT New peptides that stimulate cytokine production in T cells, useful for
 PT vaccination against human cytomegalovirus infection and for detecting
 PT immune responses to the virus.
 XX
 PS Claim 2; Page 9; 10pp; German.
 XX
 CC This invention describes novel peptides (including Immediate-Early 1
 CC Protein (IE-1) fragments) and their derivatives resulting from deletions,
 CC insertions or substitutions which are capable of inducing the secretion
 CC of interferon-gamma in CD8+ cells. The products of the invention have
 CC immunogenic activity and are useful for the production of vaccines
 CC against human cytomegaloviruses (HCMV) infections and for diagnostics for
 CC the identification of HCMV infections
 XX
 SQ Sequence 9 AA;
 Query Match 100.0%; Score 41; DS 4; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ELRRKQMY 8
 OS |||||
 DB 2 ELRRKQMY 9

RESULT 3
 AAB46142
 ID AAB46142 standard; peptide; 9 AA.
 XX
 AC AAB46142;
 XX
 DT 04-APR-2001 (first entry)
 XX
 DE HCMV type 16 antigenic peptide SEQ ID NO 10.
 XX
 KW Antigenic; vaccine; interferon-gamma; tumor necrosis factor-alpha;

KW CD8+ T cell; HCMV infection; immune response.
 XX
 OS Human cytomegalovirus.
 XX
 PN WO200075180-A2.
 XX
 PD 14-DEC-2000.
 XX
 PF 02-JUN-2000; 2000WO-DE001854.
 XX
 PR 04-JUN-1999; 99DE-01027039.
 XX
 PR 07-SEP-1999; 99DE-01043702.
 XX
 PA (KERN/) KERN F.
 XX
 PI Kern F, Volk H, Reinke P, Faulhaber N, Surel I, Khatanzas E;
 XX
 DR WPI; 2001-061700/07.
 XX
 PT New peptides that stimulate cytokine production in T cells, useful for
 PT vaccination against human cytomegalovirus infection and for detecting
 PT immune responses to the virus.
 XX
 PS Claim 1; Page 24; 28pp; German.
 XX
 CC This invention describes novel peptides (I), their encoding DNA (II) or
 CC their derivatives, that stimulate production of interferon-gamma and
 CC tumor necrosis factor-alpha in CD8+ T cells, particularly those from
 CC humans of suitable HLA type who have been immunized with human
 CC cytomegalovirus (HCMV). (I) and their derivatives are used for
 CC therapeutic or prophylactic vaccination against HCMV infection and as
 CC diagnostic reagents for detecting or quantifying the cellular immune
 CC response to HCMV (from ability to induce production of the specific
 CC cytokines in CD8+ cells), particularly in immune-deficient subjects. Also
 CC DNAs (II) that encode (I), and vectors and plasmids containing (II), are
 CC useful as pharmaceuticals
 XX
 SQ Sequence 9 AA;
 Query Match 100.0%; Score 41; DS 4; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ELRRKQMY 8
 OS |||||
 DB 2 ELRRKQMY 9

RESULT 4
 AAB46141
 ID AAB46141 standard; peptide; 9 AA.
 XX
 AC AAB46141;
 XX
 DT 04-APR-2001 (first entry)
 XX
 DE HCMV type 16 antigenic peptide SEQ ID NO 9.
 XX
 KW Antigenic; vaccine; interferon-gamma; tumor necrosis factor-alpha;
 KW CD8+ T cell; HCMV infection; immune response.
 XX
 OS Human cytomegalovirus.
 XX
 PN WO200075180-A2.
 XX
 PD 14-DEC-2000.
 XX
 PF 02-JUN-2000; 2000WO-DE001854.
 XX
 PR 04-JUN-1999; 99DE-01027039.
 XX
 PR 07-SEP-1999; 99DE-01043702.
 XX
 PA (KERN/) KERN F.

X Kern F, Volk H, Reinke P, Faulhaber N, Surel I, Khatamzas E;
X WPI: 2001-061700/07.
X
X New peptides that stimulate cytokine production in T cells, useful for
X vaccination against human cytomegalovirus infection and for detecting
X immune responses to the virus.
X
X Claim 1; Page 24; 28pp; German.
X
X This invention describes novel peptides (I), their encoding DNA (II) or
X their derivatives, that stimulate production of interferon-gamma and
X tumor necrosis factor-alpha in CD8+ T cells, particularly those from
X humans of suitable HLA type who have been immunized with human
X cytomegalovirus (hCMV). (I) and their derivatives are used for
X therapeutic or prophylactic vaccination against hCMV infection and as
X diagnostic reagents for detecting or quantifying the cellular immune
X response to hCMV (from ability to induce production of the specific
X cytokines in CD8+ cells), particularly in immune-deficient subjects. Also
X DNAs (II) that encode (I), and vectors and plasmids containing (II), are
X useful as pharmaceuticals
X
X Sequence 9 AA;
X
X Query Match 100.0%; Score 41; DB 4; Length 9;
X Best Local Similarity 100.0%; Pred. No. 1.4e+06;
X Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0
X
X Y 1 ELRRKMMY 8
X |||||
X b 1 ELRRKMMY 8
X
X RESULT 5
X ABJ38054
X .D ABJ38054 standar; peptide; 9 AA.
X
X ABJ38054;
X
X 22-MAY-2003 (first entry)
X
X Human cytomegalovirus CTL epitope peptide SEQ ID No 112.
X
X Viricide; immunostimulant; cytotoxic T-lymphocyte; CTL; epitope; antigen;
X human cytomegalovirus; hCMV; pp28; pp55; pp71; pp150; gB; gH; IE-1;
X IE-2; US2; US3; US6; US11; UL18; cytomegalovirus infection; pregnancy;
X transplantation.
X
X Human cytomegalovirus.
X
X WC2003000720-A1.
X
X 03-JAN-2003.
X
X 26-JUN-2002; 2002WO-AU000829.
X
X 26-JUN-2001; 2001AU-00005931.
X
X (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.
X
X Khanna R, Elkington RA, Walker SJ;
X
X WPI: 2003-300379/29.
X
X New human cytomegalovirus (CMV) cytotoxic T-cell epitope peptide, useful
X for diagnosing, preventing or treating CMV infection, comprises pp28,
X pp50, pp65, pp71, pp150, gB, gH, IE-1, IE-2, US2, US3, US5, US11 or UL18.
X
X Claim 1; Page 106; 308pp; English.
X
X The invention relates to a novel isolated peptide comprising one or more
X cytotoxic T-lymphocyte (CTL) epitope(s) of the same or different antigen

CC	of a human cytomegalovirus (HCMV). The HCMV antigens can comprise pp28,
CC	pp50, pp65, pp71, pp150, gB, gH, IE-1, IE-2, US2, US3, US6, US11 or US18
CC	the peptide comprises a sequence of about 9-20 contiguous amino acids of
CC	the antigen. The peptide epitopes are useful in diagnosing, preventing or
CC	treating cytomegalovirus infection in humans, and in monitoring immune
CC	responses in various clinical settings (e.g. transplantation or
CC	pregnancy). This sequence represents a human cytomegalovirus CTL epitope
CC	peptide of the invention
XX	
XX	Sequence 9 AA;
SQ	
	Query Match 100.0%; Score 41; DB 6; Length 9;
	Best Local Similarity 100.0%; Pred. No. 1.4e+06;
	Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 ELRRKQMY 8
Db	1 ELRRKQMY 8
RESULT 6	
AAB46265	
ID	AAB46265 standard; peptide, 10 AA.
XX	
AC	AAB46265;
XX	
DT	04-APR-2001 (first entry)
XX	
DE	HCMV IE-1(198-207) antigenic peptide fragment.
XX	
XW	Antigen; vaccine; human; infection; Immediate-Early 1 Protein; IE-1;
XW	interferon-gamma; CD8+ cell; immunogenic.
XX	
QS	Human cytomegalovirus.
XX	
FN	DB19927039-A1.
XX	
PD	07-DEC-2000.
XX	
PF	04-JUN-1999; 99DE-01027039.
XX	
PR	04-JUN-1999; 99DE-01027039.
PR	07-SEP-1999; 99DE-01043702.
XX	
PA	(KERN/) KERN F.
XX	
PI	Kern F, Volk H, Reinke P, Faulhaber N, Surel I, Khatamzas E;
XX	
XX	WPI; 2001-061700/07.
XX	
PR	New peptides that stimulate cytokine production in T cells, useful for
PT	vaccination against human cytomegalovirus infection and for detecting
PT	immune responses to the virus.
XX	
PS	Claim 2; Page 7; 10pp; German.
XX	
CC	This invention describes novel peptides (including Immediate-Early 1
CC	Protein (IE-1) fragments) and their derivatives resulting from deletions,
CC	insertions or substitutions which are capable of inducing the secretion
CC	of interferon-gamma in CD8+ cells. The products of the invention have
CC	immunogenic activity and are useful for the production of vaccines
CC	against human cytomegaloviruses (HCMV) infections and for diagnostics for
CC	the identification of HCMV infections
XX	
SQ	Sequence 10 AA;
	Query Match 100.0%; Score 41; DB 4; Length 10;
	Best Local Similarity 100.0%; Pred. No. 0.11;
	Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 ELRRKQMY 8
Db	2 ELRRKQMY 9

RESULT 7
AAB46146
ID AAB46146 standard; peptide; 10 AA.

XX AC AAB46146;
XX DT 04-APR-2001 (first entry)
XX DE HCMV type 16 antigenic peptide SEQ ID NO 14.
XX KW Antigenic; vaccine; interferon-gamma; tumor necrosis factor-alpha;
XX KW CD8+ T cell; hCMV infection; immune response.

XX OS Human cytomegalovirus.

XX PN WO200075180-A2.

XX PD 14-DEC-2000.

XX PF 02-JUN-2000; 2000WO-DE001854.

XX PR 04-JUN-1999; 99DE-01027039.

XX PR 07-SEP-1999; 99DE-01043702.

XX PA (KERN/) KERN F.

XX PI Kern F, Volk H, Reinke P, Faulhaber N, Surel I, Khatamzas E;

XX DR WPI; 2001-061700/07.

XX PT New peptides that stimulate cytokine production in T cells, useful for
PT vaccination against human cytomegalovirus infection and for detecting
PT immune responses to the virus.

XX PS Claim 1; Page 25; 28pp; German.

XX CC This invention describes novel peptides (I), their encoding DNA (II) or
CC their derivatives, that stimulate production of interferon-gamma and
CC tumor necrosis factor-alpha in CD8+ T cells, particularly those from
CC humans of suitable HLA type who have been immunized with human
CC cytomegalovirus (hCMV). (I) and their derivatives are used for
CC therapeutic or prophylactic vaccination against hCMV infection and as
CC diagnostic reagents for detecting or quantifying the cellular immune
CC response to hCMV (from ability to induce production of the specific
CC cytokines in CD8+ cells), particularly in immune-deficient subjects. Also
CC DNAs (II) that encode (I), and vectors and plasmids containing (II), are
CC useful as pharmaceuticals

XX SQ Sequence 10 AA;

Query Match 100.0%; Score 41; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.11; Mismatches 0; Gaps 0;
Matches 8; Conservative 0; Indels 0;

OY 1 ELRRKMMY 8
DB 2 ELRRKMMY 9

RESULT 8
AAB46135
ID AAB46135 standard; peptide; 10 AA.

XX AC AAB46135;

XX DT 04-APR-2001 (first entry)

XX DE HCMV type 16 antigenic peptide SEQ ID NO 3.

XX KW Antigenic; vaccine; interferon-gamma; tumor necrosis factor-alpha;
KW CD8+ T cell; hCMV infection; immune response.

XX OS Human cytomegalovirus.

XX PN WO200075180-A2.

XX PD 14-DEC-2000.

XX PF 02-JUN-2000; 2000WO-DE001854.

XX PR 04-JUN-1999; 99DE-01027039.

XX PR 07-SEP-1999; 99DE-01043702.

XX PA (KERN/) KERN F.

XX PI Kern F, Volk H, Reinke P, Faulhaber N, Surel I, Khatamzas E;

XX DR WPI; 2001-061700/07.

XX PT New peptides that stimulate cytokine production in T cells, useful for
PT vaccination against human cytomegalovirus infection and for detecting
PT immune responses to the virus.

XX PS Claim 1; Page 23; 28pp; German.

XX CC This invention describes novel peptides (I), their encoding DNA (II) or
CC their derivatives, that stimulate production of interferon-gamma and
CC tumor necrosis factor-alpha in CD8+ T cells, particularly those from
CC humans of suitable HLA type who have been immunized with human
CC cytomegalovirus (hCMV). (I) and their derivatives are used for
CC therapeutic or prophylactic vaccination against hCMV infection and as
CC diagnostic reagents for detecting or quantifying the cellular immune
CC response to hCMV (from ability to induce production of the specific
CC cytokines in CD8+ cells), particularly in immune-deficient subjects. Also
CC DNAs (II) that encode (I), and vectors and plasmids containing (II), are
CC useful as pharmaceuticals

XX SQ Sequence 10 AA;

Query Match 100.0%; Score 41; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.11; Mismatches 0; Gaps 0;
Matches 8; Conservative 0; Indels 0;

OY 1 ELRRKMMY 8
DB 2 ELRRKMMY 9

RESULT 9
AAB86600
ID AAB86600 standard; peptide; 15 AA.

XX AC AAB86600;

XX DT 20-NOV-2001 (first entry)

XX DE Human cytomegalovirus strain AD169 IE1 peptide fragment SEQ ID 57.

XX KW Antigen-specific stimulation; T-lymphocyte; CD8 stimulation; pp65;
KW CD4 stimulation; immuno-stimulation; IE1; lower matrix phosphoprotein.

XX OS Human cytomegalovirus.

XX PN WO200163286-A2.

XX PD 30-AUG-2001.

XX PF 17-FEB-2001; 2001WO-EP001773.

XX PR 22-FEB-2000; 2000DE-01009341.

XX PA (KERN/) KERN F.

XX PI Kern F, Volk H, Reinke P, Faulhaber N, Surel I, Khatamzas E;

XX WPI; 2001-557718/62.
 XX Stimulating CD8 or CD4 T cells for the immuno-stimulation of mammals and
 XX the determination of the response to the antigen comprises using a
 XX synthetic peptide library of the specific antigen.
 XX Example; Fig 1; 85pp; German.
 XX This sequence represents a novel method for stimulating CD8 or CD4 T
 XX cells for the immuno-stimulation of mammals and the determination of the
 XX acid sequence of the antigen (I). The method comprises (i) dividing the amino
 XX amino acids, whereby adjacent or neighbouring fragments (II) of at least 9
 XX antigen sequence; (ii) synthesizing a peptide library containing (II);
 XX and (iii) incubating the CD8 and/or CD4 T lymphocytes in a suspension
 XX comprising (II) in a single culture vessel. The method is used to immuno-
 XX stimulate T cells of mammals, particularly humans, and is also useful for
 XX diagnostic purposes to determine if a mammal, especially human, has
 XX previously immunologically responded to a specific protein, and to
 XX or in vitro immuno-stimulation of mammalian, more preferably human T
 XX lymphocytes, where the stimulated T lymphocytes are expanded and can be
 XX transduced into a patient. AAB86544-AAB86803 represent fragments of the
 XX human cytomegalovirus IE1 and pp65 lower matrix phosphoprotein which are
 XX used to illustrate the method of the invention
 XX Sequence 15 AA;
 XX
 XX Query Match 100.0%; Score 41; DB 4; Length 15;
 XX Best Local Similarity 100.0%; Pred. No. 0.16;
 XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX 1 ELRRKWMY 8
 XX |||||
 XX 7 ELRRKWMY 14
 XX
 XX RESULT 10
 XX AAB86602
 XX ID AAB86602 standard; peptide; 15 AA.
 XX
 XX AAB86602;
 XX
 XX 20-NOV-2001 (first entry)
 XX
 XX Human cytomegalovirus strain AD169 IE1 peptide fragment SEQ ID 59.
 XX
 XX Antigen-specific stimulation; T-lymphocyte; CD8 stimulation; pp65;
 XX CD4 stimulation; immuno-stimulation; IE1; lower matrix phosphoprotein.
 XX
 XX Human cytomegalovirus.
 XX
 XX WO200163286-A2.
 XX
 XX 30-AUG-2001.
 XX
 XX 17-FEB-2001; 2001WO-EP001773.
 XX
 XX 22-FEB-2000; 2000DE-01009341.
 XX
 XX (KERN/) KERN F.
 XX
 XX Kern F, Volk H, Reinke P, Faulhaber N, Surel I, Khatanzas E;
 XX WPI; 2001-557718/62.
 XX
 XX Stimulating CD8 or CD4 T cells for the immuno-stimulation of mammals and
 XX the determination of the response to the antigen comprises using a
 XX synthetic peptide library of the specific antigen.
 XX Example; Fig 1; 85pp; German.

CC This sequence represents a novel method for stimulating CD8 or CD4 T
 CC cells for the immuno-stimulation of mammals and the determination of the
 CC acid sequence of the antigen (I). The method comprises (i) dividing the amino
 CC amino acids, whereby adjacent or neighbouring fragments (II) of at least 9
 CC antigen sequence; (ii) synthesizing a peptide library containing (II);
 CC and (iii) incubating the CD8 and/or CD4 T lymphocytes in a suspension
 CC comprising (II) in a single culture vessel. The method is used to immuno-
 CC stimulate T cells of mammals, particularly humans, and is also useful for
 CC diagnostic purposes to determine if a mammal, especially human, has
 CC previously immunologically responded to a specific protein, and to
 CC or in vitro immuno-stimulation of mammalian, more preferably human T
 CC lymphocytes, where the stimulated T lymphocytes are expanded and can be
 CC transduced into a patient. AAB86544-AAB86803 represent fragments of the
 CC human cytomegalovirus IE1 and pp65 lower matrix phosphoprotein which are
 CC used to illustrate the method of the invention
 CC Sequence 15 AA;
 CC
 CC Query Match 100.0%; Score 41; DB 4; Length 15;
 CC Best Local Similarity 100.0%; Pred. No. 0.16;
 CC Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC 1 ELRRKWMY 8
 CC |||||
 CC 3 ELRRKWMY 10
 CC
 CC RESULT 11
 CC AAB46266
 CC ID AAB46266 standard; peptide; 15 AA.
 CC
 CC AAB46266;
 CC
 CC 04-APR-2001 (first entry)
 CC
 CC HCMV IE-1(199-213) antigenic peptide fragment.
 CC
 CC Antigen; vaccine; human; infection; Immediate-Early 1 Protein; IE-1;
 CC interferon-gamma; CD8+ cell; immunogenic.
 CC
 CC Human cytomegalovirus.
 CC
 CC DE19927039-A1.
 CC
 CC 07-DEC-2000.
 CC
 CC 04-JUN-1999; 99DE-01027039.
 CC
 CC 04-JUN-1999; 99DE-01027039.
 CC
 CC 07-SEP-1999; 99DE-01043702.
 CC
 CC (KERN/) KERN F.
 CC
 CC Kern F, Volk H, Reinke P, Faulhaber N, Surel I, Khatanzas E;
 CC WPI; 2001-061700/07.
 CC
 CC New peptides that stimulate cytokine production in T cells, useful for
 CC vaccination against human cytomegalovirus infection and for detecting
 CC immune responses to the virus.
 CC
 CC Claim 2; Page 8; 10pp; German.
 CC
 CC This invention describes novel peptides (including Immediate-Early 1
 CC Protein (IE-1) fragments) and their derivatives resulting from deletions,
 CC insertions or substitutions which are capable of inducing the secretion
 CC of interferon-gamma in CD8+ cells. The products of the invention have
 CC immunogenic activity and are useful for the production of vaccines
 CC against human cytomegaloviruses (HCMV) infections and for diagnostics for
 CC the identification of HCMV infections
 CC
 CC XX

SQ Sequence 15 AA;
 Query Match 100.0%; Score 41; DB 4; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.16; Mismatches 0; Indels 0; Gaps 0;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ELRRKMY 8
 Db 1 ELRRKMY 8

RESULT 12
 AAB46264
 ID AAB46264 standard; peptide; 15 AA.
 XX AC AAB46264;
 XX 04-APR-2001 (first entry)
 XX HCMV IE-1(193-207) antigenic peptide fragment.
 XX Antigen; vaccine; human; infection; Immediate-Early 1 Protein; IE-1;
 KW interferon-gamma; CD8+ cell; immunogenic.
 XX Human cytomegalovirus.
 XX DE19927039-A1.
 XX 07-DEC-2000.
 XX 04-JUN-1999; 99DE-01027039.
 XX 04-JUN-1999; 99DE-01027039.
 PR 07-SEP-1999; 99DE-01043702.
 XX (KERN/) KERN F.
 XX Kern F, Volk H, Reinke P, Faulhaber N, Surel I, Khatamzas E;
 WPI; 2001-061700/07.
 XX New peptides that stimulate cytokine production in T cells, useful for
 PT vaccination against human cytomegalovirus infection and for detecting
 PT immune responses to the virus.
 XX Claim 1; Page 23; 28pp; German.
 XX This invention describes novel peptides (I), their encoding DNA (II) or
 CC their derivatives, that stimulate production of interferon-gamma and
 CC tumor necrosis factor-alpha in CD8+ T cells, particularly those from
 CC humans of suitable HLA type who have been immunized with human
 CC cytomegalovirus (HCMV). (I) and their derivatives are used for
 CC therapeutic or prophylactic vaccination against HCMV infection and as
 CC diagnostic reagents for detecting or quantifying the cellular immune
 CC response to HCMV (from ability to induce production of the specific
 CC cytokines in CD8+ cells), particularly in immune-deficient subjects. Also
 CC DNAs (II) that encode (I), and vectors and plasmids containing (II), are
 CC useful as pharmaceuticals
 XX SQ Sequence 15 AA;
 Query Match 100.0%; Score 41; DB 4; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.16;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ELRRKMY 8
 Db 7 ELRRKMY 14

RESULT 14
 AAB46136
 ID AAB46136 standard; peptide; 15 AA.
 XX AC AAB46136;
 XX 04-APR-2001 (first entry)
 XX HCMV type 16 antigenic peptide SEQ ID NO 4.
 XX Antigenic; vaccine; interferon-gamma; tumor necrosis factor-alpha;
 KW CD8+ T cell; HCMV infection; immune response.
 XX Human cytomegalovirus.
 XX WO200075180-A2.
 XX 14-DEC-2000.
 XX 02-JUN-2000; 2000WO-DE001854.
 PF 02-JUN-2000; 2000WO-DE001854.
 XX

SQ Sequence 15 AA;
 Query Match 100.0%; Score 41; DB 4; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.16;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ELRRKMY 8
 Db 7 ELRRKMY 14

RESULT 13
 AAB46134
 ID AAB46134 standard; peptide; 15 AA.
 XX AC AAB46134;
 XX 04-APR-2001 (first entry)
 XX HCMV type 16 antigenic peptide SEQ ID NO 4.
 XX Antigenic; vaccine; interferon-gamma; tumor necrosis factor-alpha;
 KW CD8+ T cell; HCMV infection; immune response.
 XX Human cytomegalovirus.
 XX WO200075180-A2.
 XX 14-DEC-2000.
 XX 02-JUN-2000; 2000WO-DE001854.
 PF 02-JUN-2000; 2000WO-DE001854.
 XX

PR 04-JUN-1999; 93DE-01027039.
 PR 07-SEP-1999; 93DE-01043702.

XX (KERN/) KERN F.

PI Kern F, Volk H, Reinke P, Faulhaber N, Surel I, Khatamzas E;

XX WPI; 2001-061700/07.

DR New peptides that stimulate cytokine production in T cells, useful for
 PT vaccination against human cytomegalovirus infection and for detecting
 PT immune responses to the virus.

XX Claim 1; Page 23; 28pp; German.

XX This invention describes novel peptides (I), their encoding DNA (II) or
 CC their derivatives, that stimulate production of interferon-gamma and
 CC tumor necrosis factor-alpha in CD8+ T cells, particularly those from
 CC humans of suitable HLA type who have been immunized with human
 CC cytomegalovirus (HCMV). (I) and their derivatives are used for
 CC therapeutic or prophylactic vaccination against HCMV infection and as
 CC diagnostic reagents for detecting or quantifying the cellular immune
 CC response to HCMV (from ability to induce production of the specific
 CC cytokines in CD8+ cells), particularly in immune-deficient subjects. Also
 CC DNAs (II) that encode (i), and vectors and plasmids containing (II), are
 CC useful as pharmaceuticals

XX Sequence 15 AA;

Query Match 100.0%; Score 41; DB 4; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.16;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ELRRKQMY 8
 |||||
 Db 1 ELRRKQMY 8

RESULT 15

AARS58703
 ID AARS58703 standard; protein; 406 AA.

XX AARS58703;

DT 25-MAR-2003 (revised)

DT 29-MAR-1995 (first entry)

XX HCMV IE-exon-4 subunit.

XX HCMV; human cytomegalovirus; major intermediate-early protein; IE gene;
 KW adenovirus; recombinant vaccine.

XX Human cytomegalovirus.

CS WO9417810-A1.

PN 18-AUG-1994.

PD 10-FEB-1994; 94WO-US002107.

PF 12-FEB-1993; 93US-00017130.

PR (WIST-) WISTAR INST ANATOMY & BIOLOGY.

XX Plotkin SA, Ricciardi RP, Gonczol E;

XX WPI; 1994-279379/34.

XX N-PSDB; AAQ71629.

XX Adenovirus recombinant expression system useful as vaccine against HCMV -
 PT expresses immediate early exon 4 proteins.

XX Disclosure; Page 19-21; 29pp; English.

XX

CC To construct an IE-exon-4 adenovirus recombinant, PCR was used to amplify
 CC the exon-4 portion of the IE gene from purified HCMV genomic DNA (Towne
 CC strain). The primers (given in AAQ71630-31) were synthesized so as to
 CC incorporate an XbaI site, allowing insertion into the XbaI site of
 CC adenovirus vector. Primer AAQ71630 was also modified so that an ATG codon
 CC was inserted at the first amino acid position of exon-4. (Updated on 25-
 CC MAR-2003 to correct PN field.)

XX Sequence 406 AA;

Query Match 100.0%; Score 41; DB 2; Length 406;
 Best Local Similarity 100.0%; Pred. No. 3.7;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ELRRKQMY 8
 |||||
 Db 114 ELRRKQMY 121

Search completed: May 21, 2004, 12:47:27
 Job time : 36.1304 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

2M protein - protein search, using sw model

Run on: May 21, 2004, 12:45:39 ; Search time 11.1304 Seconds
(without alignments)
37.106 Million cell updates/sec

Title: US-09-980-058-9_COPY_1_8

Perfect score: 41

Sequence: 1 ELRKQMY 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA:**
- 1: /cgn2_6/ptodata/2/iaa/5A COMB.pcp.*
 - 2: /cgn2_6/ptodata/2/iaa/5B COMB.pcp.*
 - 3: /cgn2_6/ptodata/2/iaa/6A COMB.pcp.*
 - 4: /cgn2_6/ptodata/2/iaa/6B COMB.pcp.*
 - 5: /cgn2_6/ptodata/2/iaa/PCUS COMB.pcp.*
 - 6: /cgn2_6/ptodata/2/iaa/backfiles.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	406	4	US-09-171-699-4
2	41	100.0	406	5	PCT-US94-02107-2
3	32	78.0	262	4	US-09-543-681A-5824
4	31	75.6	93	4	US-09-581-122A-29
5	31	75.6	1024	4	US-09-562-737-42
6	30	73.2	308	4	US-09-711-164-392
7	29	70.7	440	1	US-08-459-100A-2
8	29	70.7	440	5	PCT-US94-09589-2
9	29	70.7	462	4	US-09-198-452A-668
10	29	70.7	703	4	US-08-311-731A-152
11	28	68.3	196	4	US-09-247-155-163
12	28	68.3	278	4	US-09-107-532A-5114
13	28	68.3	345	4	US-09-543-681A-5409
14	28	68.3	369	3	US-09-108-020-40
15	28	68.3	554	4	US-08-809-802-8
16	28	68.3	585	4	US-09-462-951B-1
17	28	68.3	821	1	US-09-377-465A-2
18	28	68.3	1333	4	US-09-347-878-20
19	27	65.9	54	3	US-08-630-916A-124
20	27	65.9	71	4	US-09-134-000C-4679
21	27	65.9	87	4	US-09-328-352-7319
22	27	65.9	137	4	US-09-732-210-686
23	27	65.9	141	4	US-09-732-210-758
24	27	65.9	176	1	US-07-953-230A-9
25	27	65.9	178	4	US-09-489-039A-10005
26	27	65.9	183	4	US-09-489-039A-7793
27	27	65.9	198	3	US-08-842-306B-2

Sequence 2, Appli
Sequence 2, Appli
Sequence 11, Appli
Sequence 11, Appli
Sequence 11, Appli
Sequence 11, Appli
Sequence 23899, A
Sequence 236, App
Sequence 13384, A
Sequence 2275, Ap
Sequence 5382, Ap
Sequence 6605, Ap
Sequence 4237, Ap
Sequence 4, Appli
Sequence 31, Appli
Sequence 3, Appli
Sequence 2, Appli
Sequence 5167, Ap

ALIGNMENTS

RESULT 1

US-09-171-699-4

; Sequence 4, Application US/09171699

; Patent No. 6448389

; GENERAL INFORMATION:

; APPLICANT: The Wistar Institute of, Anatomy & Biology

; Gonczol, Rva

; Berencsi, Klara

; Kari, Csaba

TITLE OF INVENTION: No. 6448389el Cytomegalovirus DNA Constructs and

USes Therefor

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: Howson and Howson

STREET: Spring House Corporate Center, P.O. Box 457

CITY: Spring House

STATE: Pennsylvania

COUNTRY: USA

ZIP: 19477

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/171,699

FILING DATE: 19-Jan-1999

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/015,717

FILING DATE: 23-APR-1996

ATTORNEY/AGENT INFORMATION:

NAME: Kodroff, Cathy A.

REGISTRATION NUMBER: 33,980

REFERENCE/DOCKET NUMBER: WST66APCT

TELECOMMUNICATION INFORMATION:

TELEPHONE: 215-540-9200

TELEFAX: 215-540-5818

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 406 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-09-171-699-4

Query Match 100.0%; Score 41; DB 4; Length 406;

Best Local Similarity 100.0%; Pred. No. 1.9;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ELRRKMY 8
Db 114 ELRRKMY 121

RESULT 2
PCT-US94-02107-2
; Sequence 2, Application PC/TUS9402107
; GENERAL INFORMATION:
; APPLICANT: The Wistar Institute of Anatomy and Biology
; TITLE OF INVENTION: Recombinant Cytomegalovirus Vaccine
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, PO Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/02107
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/017,130
; FILING DATE: 12-FEB-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Sak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: WST6BPCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 406 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US94-02107-2

Query Match 100.0%; Score 41; DB 5; Length 406;
Best Local Similarity 100.0%; Pred. No. 1.9; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0

Qy 1 ELRRKMY 8
Db 114 ELRRKMY 121

RESULT 3
US-09-543-681A-5824
; Sequence 5824, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5824
; LENGTH: 262
; TYPE: PRT

; ORGANISM: Proteus mirabilis
US-09-543-681A-5824

Query Match 78.0%; Score 32; DB 4; Length 262;
Best Local Similarity 62.5%; Pred. No. 64; Indels 0; Gaps 0;
Matches 5; Conservative 3; Mismatches 0

Qy 1 ELRRKMY 8
Db 16 ELRRKMY 23

RESULT 4
US-09-381-122A-29
; Sequence 29, Application US/09381122A
; Patent No. 6551594
; GENERAL INFORMATION:
; APPLICANT: Milligen van, Florine J.
; APPLICANT: Bernadus, J. Johannes B.W.J
; APPLICANT: Cornelissen, Bernard A.
; TITLE OF INVENTION: An ex vivo animal or challenge model as method to
; TITLE OF INVENTION: measure protective immunity directed against parasites
; TITLE OF INVENTION: and vaccines shown to be protective in said method
; FILE REFERENCE: 2183-4188US
; CURRENT APPLICATION NUMBER: US/09/381,122A
; CURRENT FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: PCT/NL98/00146
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: EP 97200730.6
; PRIOR FILING DATE: 1997-03-11
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 29
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Fasciola hepatica
; NAME/KEY: CHAIN
; LOCATION: (1)..(93)
; OTHER INFORMATION: /No. 6551594e="cathepsin L proregion"
US-09-381-122A-29

Query Match 75.6%; Score 31; DB 4; Length 93;
Best Local Similarity 62.5%; Pred. No. 41; Indels 0; Gaps 0;
Matches 5; Conservative 3; Mismatches 0

Qy 1 ELRRKMY 8
Db 23 ELRRKMY 30

RESULT 5
US-09-562-737-42
; Sequence 42, Application US/09562737
; Patent No. 6428967
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; APPLICANT: Gotthardt, Michael
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/09/562,737
; CURRENT FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 42
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-562-737-42

Query Match 75.6%; Score 31; DB 4; Length 1024;
Best Local Similarity 75.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ELRRKMY 8

Db 192 ELRRKFF 199

RESULT 6

US-09-711-164-392
Sequence 392, Application US/09711164
Patent No. 6589738
GENERAL INFORMATION:
APPLICANT: Forsyth, R. Allyn
APPLICANT: Ohlsen, Karl
TITLE OF INVENTION: GENES ESSENTIAL FOR MICROBIAL PROLIFERATION AND ANTISENSE THERAPY
FILE REFERENCE: ELITRA.008A
CURRENT APPLICATION NUMBER: US/09/711.164
CURRENT FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: US 60/164415
PRIOR FILING DATE: 1999-11-9
NUMBER OF SEQ ID NOS: 469
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 392
LENGTH: 308
TYPE: PRT
ORGANISM: Escherichia coli
US-09-711-164-392

Query Match 73.2%; Score 30; DB 4; Length 308;
Best Local Similarity 75.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ELRRKMY 8

Db 72 ELRRMGY 79

RESULT 7

US-08-459-100A-2
Sequence 2, Application US/08459100A
Patent No. 5654172
GENERAL INFORMATION:
APPLICANT: Li, ET AL.
TITLE OF INVENTION: GABAA Receptor Epsilon Subunit
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
ADDRESSEE: CECCHI, STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
FILING DATE: June 2, 1995
APPLICATION NUMBER: US/08/459.100A
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/09589
FILING DATE: 26 AUG 1994
ATTORNEY/AGENT INFORMATION:
NAME: MULLINS, J.G.
REGISTRATION NUMBER: 33,073
REFERENCE/DOCKET NUMBER: 325800-325

TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 440 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
US-08-459-100A-2

Query Match 70.7%; Score 29; DB 1; Length 440;
Best Local Similarity 62.5%; Pred. No. 3.7e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ELRRKMY 8

Db 237 ELRRNVLY 244

RESULT 8

PCT-US94-09589-2
Sequence 2, Application PC/TUS9409589
GENERAL INFORMATION:
APPLICANT: KIRKNESS, ET AL.
TITLE OF INVENTION: GABAA Receptor Epsilon Subunit
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
ADDRESSEE: CECCHI, STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/09589
FILING DATE: Submitted herewith
CLASSIFICATION:
PRIOR APPLICATION NUMBER:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: FERRARO, GREGORY D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-188
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 440 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
PCT-US94-09589-2

Query Match 70.7%; Score 29; DB 5; Length 440;
Best Local Similarity 62.5%; Pred. No. 3.7e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ELRRKMY 8

Db 237 ELRRNVLY 244

```
RESULT 9
US-09-198-452A-668
; Sequence 668, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Grifffais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 668
; LENGTH: 462
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-668

Query Match      70.7%; Score 29; DB 4; Length 462;
Best Local Similarity 62.5%; Pred. No. 3.8e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 ELRRKMY 8
      |||:|:|
Db      17 ELRQKLFY 24

RESULT 10
US-08-311-731A-152
; Sequence 152, Application US/08311731A
; Patent No. 6583266
; GENERAL INFORMATION:
; APPLICANT: SMITH, DOUGLAS
; APPLICANT: MAO, JEN-I
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 411
; CORRESPONDENCE ADDRESS:
; ADDRESSES: WOLF, GREENFIELD & SACKS, P.C.
; STREET: 600 ATLANTIC AVENUE
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311,731A
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: GATES, EDWARD R.
; REGISTRATION NUMBER: 31,616
; REFERENCE/DOCKET NUMBER: C0044/7125
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/720-3500
; TELEFAX: 617/720-2441
; INFORMATION FOR SEQ ID NO: 152:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 703 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Mycobacterium leprae
US-08-311-731A-152
```

```
Query Match      70.7%; Score 29; DB 4; Length 703;
Best Local Similarity 71.4%; Pred. No. 5.5e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      2 LRKMY 8
      |||:|:|
Db      76 LRQMY 82

RESULT 11
US-09-247-155-163
; Sequence 163, Application US/09247155A
; Patent No. 6312922
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Duclert, Aymeric
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Complementary DNAs
; FILE REFERENCE: GENSET.021A
; CURRENT APPLICATION NUMBER: US/09/247,155A
; CURRENT FILING DATE: 1999-02-09
; EARLIER APPLICATION NUMBER: 60/074,121
; EARLIER FILING DATE: 1998-02-09
; EARLIER APPLICATION NUMBER: 60/081,563
; EARLIER FILING DATE: 1998-04-13
; EARLIER APPLICATION NUMBER: 60/096,116
; EARLIER FILING DATE: 1998-08-10
; EARLIER APPLICATION NUMBER: 60/099,273
; EARLIER FILING DATE: 1998-10-04
; NUMBER OF SEQ ID NOS: 182
; SOFTWARE: Patent.pm
; SEQ ID NO 163
; LENGTH: 196
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -34...-1
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 81,84,87,131,135,143,156
; OTHER INFORMATION: Xaa = any one of the twenty amino acids
US-09-247-155-163

Query Match      68.3%; Score 28; DB 4; Length 196;
Best Local Similarity 83.3%; Pred. No. 2.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      3 RRKMY 8
      |||:|:|
Db      145 RRQMF 150

RESULT 12
US-09-107-532A-5114
; Sequence 5114, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; CORRESPONDENCE ADDRESS: ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; ADDRESS: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
```

SOFTWARE: ASCII
CURRENT APPLICATION DATA: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneka
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 5114:
SEQUENCE CHARACTERISTICS:
LENGTH: 278 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc_feature
LOCATION: (B) LOCATION 1...278
SEQUENCE DESCRIPTION: SEQ ID NO: 5114:
JS-09-107-532A-5114

Query Match 68.3%; Score 28; DB 4; Length 278;
Best Local Similarity 83.3%; Pred. No. 3.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Y 3 LRKXMY 8
|||:
b 6 RKIMY 11

RESULT 13
US-09-543-681A-5409
Sequence 5409, Application US/09543681A
Patent No. 6605709
GENERAL INFORMATION:
APPLICANT: GARY BRETON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
FILE REFERENCE: 2709.1002-001
CURRENT APPLICATION NUMBER: US/09/543,681A
CURRENT FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/128,706
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 8344
SEQ ID NO 5409
LENGTH: 345
TYPE: PRT
ORGANISM: Proteus mirabilis
US-09-543-681A-5409

Query Match 68.3%; Score 28; DB 4; Length 345;
Best Local Similarity 71.4%; Pred. No. 4.6e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Y 2 LRKXMY 8
|||:
b 7 LRKXMY 13

RESULT 14
US-09-108-020-40
Sequence 40, Application US/09108020A
Patent No. 6143561

GENERAL INFORMATION:
APPLICANT: Randall, Douglas D.
APPLICANT: Johnston, Mark L.
APPLICANT: Miernyk, Jan A.
APPLICANT: Luethy, Michael H.
APPLICANT: Mooney, Brian P.
TITLE OF INVENTION: USE OF DNA ENCODING PLASTID PYRUVATE DEHYDROGENASE AND
BRANCHED CHAIN OXOACID DEHYDROGENASE COMPONENTS TO
TITLE OF INVENTION: ENHANCE POLYHYDROXYALKANOATE BIOSYNTHESIS IN PLANTS
FILE REFERENCE: UMO 1482
CURRENT APPLICATION NUMBER: US/09/108,020A
CURRENT FILING DATE: 1998-06-30
EARLIER APPLICATION NUMBER: 60/051,291
EARLIER FILING DATE: 1997-06-30
EARLIER APPLICATION NUMBER: 60/055,255
EARLIER FILING DATE: 1997-08-01
EARLIER APPLICATION NUMBER: 60/076,544
EARLIER FILING DATE: 1998-03-02
NUMBER OF SEQ ID NOS: 54
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 40
LENGTH: 369
TYPE: PRT
ORGANISM: B. subtilis
US-09-108-020-40

Query Match 68.3%; Score 28; DB 3; Length 369;
Best Local Similarity 62.5%; Pred. No. 4.9e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Y 1 LRKXMY 8
|||:
b 49 ELGRMY 56

RESULT 15
US-08-809-802-8
Sequence 8, Application US/08809802
Patent No. 6455276
GENERAL INFORMATION:
APPLICANT: Le Bourdelles, Beatrice
APPLICANT: Whiting, Paul John
TITLE OF INVENTION: HUMAN ALPHA 4 RECEPTOR SUBUNIT
TITLE OF INVENTION: OF THE GABA-A RECEPTOR
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/809,802
FILING DATE: 19-JUN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB95/02323
FILING DATE: 29-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: Yablonsky, Michael D.
REGISTRATION NUMBER: 40,407
REFERENCE/DOCKET NUMBER: T1292
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-4678
TELEFAX: 732-594-4720
TELEX:
INFORMATION FOR SEQ ID NO: 8:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 554 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-809-802-8

Query Match 68.3%; Score 28; DB 4; Length 554;
Best Local Similarity 85.7%; Pred. NO. 6.9e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LRRKMY 8
| | | | |
Db 252 LRRKMY 258

Search completed: May 21, 2004, 12:50:51
Job time : 12.1304 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 21, 2004, 12:47:34 ; Search time 26.7826 Seconds
(without alignments)
83.314 Million cell updates/sec

Title: US-09-980-058-9_COPY_1_8
Perfect score: 41
Sequence: 1 ELRRKMY 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1149313 seqs, 278921704 residues

Total number of hits satisfying chosen parameters: 1149313

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pcp.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pcp.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pcp.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pcp.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pcp.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pcp.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pcp.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pcp.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pcp.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pcp.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pcp.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pcp.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pcp.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pcp.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pcp.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pcp.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pcp.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

result No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	406	14	US-10-223-538-4
2	34	82.9	144	9	US-09-924-358-25
3	34	82.9	144	15	US-10-410-764-25
4	33	80.5	106	14	US-10-029-386-28995
5	32	78.0	105	12	US-10-424-599-180529
6	32	78.0	256	12	US-10-424-599-276089
7	32	78.0	331	12	US-10-425-114-49481
8	32	78.0	510	12	US-10-424-599-194732
9	32	78.0	516	12	US-10-424-599-194730
10	32	78.0	523	12	US-10-425-114-58445
11	31	75.6	93	12	US-10-382-478-29
12	31	75.6	426	15	US-10-369-493-6827
13	31	75.6	495	12	US-10-425-114-70884
14	31	75.6	803	15	US-10-320-797-3312
15	31	75.6	1024	14	US-10-211-962-42

Sequence 144412, A
Sequence 38669, A
Sequence 220780, A
Sequence 10212, A
Sequence 56603, A
Sequence 392, App
Sequence 893, App
Sequence 72901, A
Sequence 9862, Ap
Sequence 133801, A
Sequence 42052, A
Sequence 37456, A
Sequence 1222, Ap
Sequence 149859, A
Sequence 38479, A
Sequence 265382, A
Sequence 70440, A
Sequence 62004, A
Sequence 41, Appl
Sequence 41, Appl
Sequence 41, Appl
Sequence 51338, A
Sequence 40, Appl
Sequence 40, Appl
Sequence 40, Appl
Sequence 40, Appl
Sequence 44616, A
Sequence 53364, A
Sequence 33298, A

US-10-424-599-144412
US-10-425-114-38669
US-10-424-599-220780
US-09-815-242-10212
US-10-282-122A-56603
US-10-287-274-392
US-10-369-493-893
US-10-282-122A-72901
US-10-369-493-9862
US-10-424-599-153801
US-10-425-114-42052
US-10-425-114-37456
US-10-094-749-1922
US-10-424-599-149859
US-10-425-114-38479
US-10-424-599-265382
US-10-425-114-70440
US-10-425-114-62004
US-09-778-320-41
US-09-910-689-41
US-09-510-662A-41
US-10-010-742-41
US-10-425-114-51338
US-09-778-320-40
US-09-910-689-40
US-09-510-662A-40
US-10-010-742-40
US-10-425-114-44616
US-10-282-122A-53364
US-10-369-493-23298

ALIGNMENTS

RESULT 1

US-10-223-538-4
; Sequence 4, Application US/10223538
; Publication No. US20030120060A1

GENERAL INFORMATION:

APPLICANT: The Wistar Institute of, Anatomy & Biology

Gonczol, Eva
Berencsi, Klara
Kari, Csaba

TITLE OF INVENTION:

No. US20030120060A1el Cytomegalovirus DNA Constructs and

NUMBER OF SEQUENCES:

10

CORRESPONDENCE ADDRESS:

ADDRESSEE: Howson and Howson

STREET: Spring House Corporate Center, P.O. Box 457

CITY: Spring House

STATE: Pennsylvania

COUNTRY: USA

ZIP: 19477

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/223,538

FILING DATE: 19-Aug-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/171,699

FILING DATE: 19-Jan-1999

APPLICATION NUMBER: US 60/015,717

FILING DATE: 23-APR-1996

ATTORNEY/AGENT INFORMATION:

NAME: Kodroff, Cathy A.

REGISTRATION NUMBER: 33,980

REFERENCE/DOCKET NUMBER: WST66APCT

TELECOMMUNICATION INFORMATION:

TELEPHONE: 215-540-9200
 TELEFAX: 215-540-5818
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 406 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 4:
 US-10-223-538-4

Query Match 100.0%; Score 41; DB 14; Length 406;
 Best Local Similarity 100.0%; Pred. No. 3.5;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ELRRKMY 8
 |||||
 Db 114 ELRRKMY 121

RESULT 2
 US-09-924-358-25
 Sequence 25, Application US/09924358
 Patent No. US20020107376A1
 GENERAL INFORMATION:
 APPLICANT: Millennium Pharmaceuticals, Inc.
 APPLICANT: Meyers, Rachel
 APPLICANT: MacBeth, Kyle
 TITLE OF INVENTION: 26199, 33530, 33949, 47148, 50226, AND
 TITLE OF INVENTION: 58764,
 FILE REFERENCE: 38155-20034.00
 CURRENT APPLICATION NUMBER: US/09/924,358
 CURRENT FILING DATE: 2001-08-06
 PRIOR APPLICATION NUMBER: US 60/229,300
 PRIOR FILING DATE: 2000-09-01
 NUMBER OF SEQ ID NOS: 52
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 25
 LENGTH: 144
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Consensus amino acid
 US-09-924-358-25

Query Match 82.9%; Score 34; DB 9; Length 144;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ELRRKMY 7
 |||||
 Db 111 ELRRKMY 117

RESULT 3
 US-10-410-764-25
 Sequence 25, Application US/10410764
 Publication No. US20040005664A1
 GENERAL INFORMATION:
 APPLICANT: Millennium Pharmaceuticals, Inc.
 APPLICANT: Meyers, Rachel E.
 APPLICANT: MacBeth, Kyle E.
 APPLICANT: Curtis, Rory A.J.
 APPLICANT: Rudolph-Owen, Laura A.
 APPLICANT: Weich, Nadine S.
 APPLICANT: Olandt, Peter J.
 APPLICANT: Tsai, Fong-Ying
 APPLICANT: Kapeller-Liberman, Rosana
 APPLICANT: Carroll, Joseph M.
 TITLE OF INVENTION: 26199, 33530, 33949, 47148, 50226,
 TITLE OF INVENTION: 58764, 62113, 32144, 32235, 23565, 13305, 14911, 86216,
 TITLE OF INVENTION: 25206 AND 8843 MOLECULES AND USES THEREFOR

FILE REFERENCE: MPI03-0520NMIM
 CURRENT APPLICATION NUMBER: US/10/410,764
 CURRENT FILING DATE: 2003-04-10
 PRIOR APPLICATION NUMBER: US 09/924,358
 PRIOR FILING DATE: 2001-08-06
 PRIOR APPLICATION NUMBER: US 60/229,300
 PRIOR FILING DATE: 2000-09-01
 PRIOR APPLICATION NUMBER: US 10/350,553
 PRIOR FILING DATE: 2003-01-24
 PRIOR APPLICATION NUMBER: US 60/351,572
 PRIOR FILING DATE: 2002-01-24
 PRIOR APPLICATION NUMBER: US 09/966,614
 PRIOR FILING DATE: 2001-09-27
 PRIOR APPLICATION NUMBER: US 60/238,054
 PRIOR FILING DATE: 2000-10-05
 PRIOR APPLICATION NUMBER: US 10/281,094
 PRIOR FILING DATE: 2002-10-25
 PRIOR APPLICATION NUMBER: US 60/347,815
 PRIOR FILING DATE: 2001-10-29
 PRIOR APPLICATION NUMBER: US 10/076,535
 PRIOR FILING DATE: 2002-02-15
 PRIOR APPLICATION NUMBER: US 60/269,440
 PRIOR FILING DATE: 2001-02-16
 Remaining Prior Application data removed - See File Wrapper or PALM.
 NUMBER OF SEQ ID NOS: 136
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 25
 LENGTH: 144
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Amino acid consensus sequence
 US-10-410-764-25

Query Match 82.9%; Score 34; DB 15; Length 144;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ELRRKMY 7
 |||||
 Db 111 ELRRKMY 117

RESULT 4
 US-10-029-386-28995
 Sequence 28995, Application US/10029386
 Publication No. US20030194704A1
 GENERAL INFORMATION:
 APPLICANT: Penn, Sharron G.
 APPLICANT: Rank, David R.
 APPLICANT: Hantzel, David K.
 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
 FILE REFERENCE: ABOMICA-X-2
 CURRENT APPLICATION NUMBER: US/10/029,386
 CURRENT FILING DATE: 2001-12-20
 NUMBER OF SEQ ID NOS: 34288
 SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
 SEQ ID NO 28995
 LENGTH: 106
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: MAP TO CHR22.1
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.86
 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.89
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.87
 OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.58
 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1
 OTHER INFORMATION: SWISSPROT HIT: P30754, EVALUATE 8.30e-01
 US-10-029-386-28995

Query Match 80.5%; Score 33; DB 14; Length 106;
Best Local Similarity 75.0%; Pred. No. 39;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ELRRKMY 8
|||||
DB 91 ELRRSMLY 98

RESULT 5

US-10-424-599-180529
; Sequence 180529, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 180529
; LENGTH: 105
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_134031C.1.pap
US-10-424-599-180529

Query Match 78.0%; Score 32; DB 12; Length 105;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RRKMY 8
|||||
DB 37 RRKMY 42

RESULT 6

US-10-424-599-276089
; Sequence 276089, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 276089
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_91328C.1.pap
US-10-424-599-276089

Query Match 78.0%; Score 32; DB 12; Length 256;
Best Local Similarity 75.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ELRRKMY 8
|||||
DB 183 ELRRKMY 190

RESULT 7

US-10-425-114-49481
; Sequence 49481, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53113)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 49481
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 701205658_FELI.pap
US-10-425-114-49481

Query Match 78.0%; Score 32; DB 12; Length 331;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RRKMY 8
|||||
DB 306 RRKMY 311

RESULT 8

US-10-424-599-194732
; Sequence 194732, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 194732
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1) (510)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_17870C.1.pap
US-10-424-599-194732

Query Match 78.0%; Score 32; DB 12; Length 510;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RRKMY 8
|||||
DB 302 RRKMY 307

RESULT 9

US-10-424-599-194730
; Sequence 194730, Application US/10424599

```
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; PRIOR FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 194730
; LENGTH: 516
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(516)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_17869C.1.pap
US-10-424-599-194730

Query Match      78.0%; Score 32; DB 12; Length 516;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3 RRRKMY 8
Db      309 RRRKMY 314

RESULT 10
US-10-425-114-58445
; Sequence 58445, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; PRIOR FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 58445
; LENGTH: 523
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-GMFLMINSOX070C11_FLI.pap
US-10-425-114-58445

Query Match      78.0%; Score 32; DB 12; Length 523;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3 RRRKMY 8
Db      319 RRRKMY 324

RESULT 11
US-10-382-479-29
; Sequence 29, Application US/10382479
; Publication No. US20030224007A1
; GENERAL INFORMATION:
; APPLICANT: Milligen van, Florine J.
```

```
; APPLICANT: Bernardus, J.
; APPLICANT: Cornelissen, Johannes B.W.J
; APPLICANT: Bokhout, Bernard A.
; TITLE OF INVENTION: An ex vivo animal or challenge model as method to
; FILE REFERENCE: 2183-4198US
; CURRENT APPLICATION NUMBER: US/10/382,479
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: US/09/381,122A
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: PCT/NL98/00146
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: EP 97200730.6
; PRIOR FILING DATE: 1997-03-11
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patentin Ver. 3.0
; SEQ ID NO 29
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Fasciola hepatica
; FEATURE:
; NAME/KEY: CHAIN
; LOCATION: (1)..(93)
; OTHER INFORMATION: /Note="cathepsin L proregion"
US-10-382-479-29

Query Match      75.6%; Score 31; DB 12; Length 93;
Best Local Similarity 62.5%; Pred. No. 85;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ELRRKMY 8
Db      23 ELRRKMY 30

RESULT 12
US-10-369-493-6827
; Sequence 6827, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 6827
; LENGTH: 426
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-6827

Query Match      75.6%; Score 31; DB 15; Length 426;
Best Local Similarity 75.0%; Pred. No. 3.8e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1 ELRRKMY 8
Db      229 ELRRKMY 236

RESULT 13
US-10-425-114-70884
; Sequence 70884, Application US/10425114
; Publication No. US20040034888A1
```

GENERAL INFORMATION:
APPLICANT: Herz, Joachim
APPLICANT: Gotthardt, Michael
TITLE OF INVENTION: LDL Receptor Signaling Pathways
FILE REFERENCE: UTSW0708
CURRENT APPLICATION NUMBER: US/10/211,962
CURRENT FILING DATE: 2002-09-01
PRIOR APPLICATION NUMBER: US/09/562,737
PRIOR FILING DATE: 2000-05-01
NUMBER OF SEQ ID NOS: 132
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 42
LENGTH: 1024
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-211-962-42

Query Match 75.6%; Score 31; DB 14; Length 1024;
Best Local Similarity 75.0%; Pred. No. 8.8e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ELRRKMMY 8
| | | | |
Db 192 ELRRKPPY 199

Search completed: May 21, 2004, 12:52:20
Job time : 26.7826 secs

GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 70884
LENGTH: 495
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: LIB4172-034-D4_FLI.pep
S-10-425-114-70884

Query Match 75.6%; Score 31; DB 12; Length 495;
Best Local Similarity 85.7%; Pred. No. 4.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

y 1 ELRRROM 7
| | | | |
b 6 EVRRROM 12

RESULT 14
IS-10-320-797-3312
Sequence 3312, Application US/10320797
Publication No. US20040014955A1
GENERAL INFORMATION:
APPLICANT: Eroshkin, Alexey M.
APPLICANT: Zamudio, Carlos
TITLE OF INVENTION: IDENTIFICATION OF ESSENTIAL GENES OF CRYPTOCOCCUS NEOFORMANS AND
TITLE OF INVENTION: METHODS OF USE
FILE REFERENCE: 10182-021-999
CURRENT APPLICATION NUMBER: US/10/320,797
CURRENT FILING DATE: 2002-12-16
PRIOR APPLICATION NUMBER: 60/341,261
PRIOR FILING DATE: 2001-12-17
NUMBER OF SEQ ID NOS: 3361
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3312
LENGTH: 803
TYPE: PRT
ORGANISM: Cryptococcus neoformans
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (12)..(12)
OTHER INFORMATION: Xaa = any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (462)..(462)
OTHER INFORMATION: Xaa = any amino acid
IS-10-320-797-3312

Query Match 75.6%; Score 31; DB 15; Length 803;
Best Local Similarity 75.0%; Pred. No. 7e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ELRRKMMY 8
| | | | |
Db 89 ELRLMMY 96

RESULT 15
US-10-211-962-42
Sequence 42, Application US/10211962
Publication No. US20030082640A1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

M protein - protein search, using sw model

on on: May 21, 2004, 12:44:59 ; Search time 9.3913 Seconds
(without alignments)

81.941 Million cell updates/sec

itle: US-09-980-058-9_COPY_1_8

effect score: 41

equences: 1 ELRRKXMY 8

coring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

earched: 283366 seqs, 96191526 residues

otal number of hits satisfying chosen parameters: 283366

inimum DB seq length: 0

Maximum DB seq length: 2000000000

ost-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

atabase :

1: PIR 78.*

2: PIR1.*

3: PIR2.*

4: PIR3.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

result	No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	491	1	EDBEIC	immediate-early pr
2	41	100.0	491	1	EDBEIM5	immediate-early pr
3	32	78.0	475	2	C72538	probable cysteinyl
4	31	75.6	102	2	B72571	probable ribosomal
5	31	75.6	305	2	D69024	hypothetical prote
6	31	75.6	426	2	T16406	hypothetical prote
7	31	75.6	756	2	G84866	hypothetical prote
8	31	75.6	907	2	E96692	probable wall-asso
9	31	75.6	1331	1	XORTDH	xanthine dehydrog
10	31	75.6	1335	1	XOMSDH	xanthine dehydrog
11	31	75.6	2156	1	RRVUNE	genome polyprotein
12	30	73.2	156	2	D72221	transcription elon
13	30	73.2	256	2	D81181	conserved hypothet
14	30	73.2	256	2	C81925	probable sec-indop
15	30	73.2	263	2	A13081	conserved hypothet
16	30	73.2	263	2	G98204	lactam utilization
17	30	73.2	308	2	H91005	hypothetical prote
18	30	73.2	308	2	A85850	hypothetical prote
19	30	73.2	308	2	H64980	hypothetical ABC t
20	30	73.2	315	2	A10777	ABC transporter AT
21	30	73.2	328	2	T40910	probable tyrosine
22	30	73.2	331	2	H83693	hypothetical prote
23	30	73.2	371	1	S59442	DNA repair protein
24	30	73.2	776	2	AH1147	preprotein translo
25	30	73.2	776	2	AH1506	preprotein translo
26	30	73.2	1091	2	T34247	hypothetical prote
27	30	73.2	1107	2	T34246	hypothetical prote
28	30	73.2	1216	2	T34101	hypothetical prote
29	30	73.2	1740	2	T43773	hypothetical prote

ALIGNMENTS

RESULT 1

EDBEIC

immediate-early protein - human cytomegalovirus (strain Towne)

C/Species: human cytomegalovirus, human herpesvirus 5

C/Date: 20-Sep-1984 #sequence_revision 20-Sep-1984 #text_change 16-Jul-1999

C/Accession: A03722; S34420

R/Stenberg, R.M.; Thomsen, D.R.; Stinski, M.F.

J. Virol. 49, 190-199, 1984

A/Title: Structural analysis of the major immediate early gene of human cytomegalovirus

A/Reference number: A03722; MUID:84090395; PMID:6317889

A/Accession: A03722

A/Molecule type: DNA

A/Residues: 1-491 <STE>

A/Note: the authors translated the codons GAG, TCC, and CGT for residues 245, 313, and 3

R/Chapman, B.S.; Thayer, R.M.; Vincent, K.A.; Haigwood, N.L.

Nucleic Acids Res. 19, 3979-3986, 1991

A/Title: Effect of intron A from human cytomegalovirus (Towne) immediate-early gene on l

A/Reference number: S34420; MUID:91319560; PMID:1650459

A/Accession: S34420

A/Status: translation not shown

A/Molecule type: DNA

A/Residues: 1-48 <CHA>

A/Cross-references: EMBL:M60321; NID:g330624; PIDN:AAA45982.1; PID:g330625

C/Genetics:

A/Introns: 24/2

C/Superfamily: cytomegalovirus immediate-early protein

Query Match 100.0%; Score 41; DB 1; Length 491;

Best Local Similarity 100.0%; Pred. NO. 0.61;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ELRRKXMY 8

DB 199 ELRRKXMY 206

RESULT 2

EDBEIM5

immediate-early protein - human cytomegalovirus (strain AD169)

N/Alternate names: UI123 protein

C/Species: human cytomegalovirus, human herpesvirus 5

A/Note: host Homo sapiens (man)

C/Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 16-Jun-2000

C/Accession: S09890

R/Chen, M.S.; Bankier, A.T.; Beck, S.; Bohni, R.; Brown, C.M.; Cerny, R.; Horsnell, T.;

M.; Barrell, B.G.

Curr. Top. Microbiol. Immunol. 154, 125-169, 1990

A/Title: Analysis of the protein-coding content of the sequence of human cytomegalovirus

A/Reference number: S09749; MUID:90269039; PMID:2161319

A/Accession: S09890

A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

actin-depolymerizi
transcription regu
conserved hypothet
conserved hypothet
hypothetical prote
hypothetical prote
beta-lactamase [m
glycine betaine/ca
choline ABC transp
conserved hypothet
hypothetical prote
translocase [impor
translocase - Chla
preprotein translo
hypothetical prote
phenylalanine-tRNA

30 29 70.7 130 2 T05767
31 29 70.7 169 2 AE1166
32 29 70.7 174 2 C69539
33 29 70.7 192 2 C63174
34 29 70.7 218 2 A84688
35 29 70.7 242 2 F84788
36 29 70.7 260 2 A96923
37 29 70.7 380 2 C69670
38 29 70.7 381 2 G69669
39 29 70.7 406 2 A72382
40 29 70.7 440 2 T32151
41 29 70.7 457 2 B85569
42 29 70.7 457 2 G72053
43 29 70.7 462 2 B81613
44 29 70.7 473 2 S64623
45 29 70.7 556 2 B71111

```

;Residues: 1-491 <CHE>
;Cross-references: EMBL:X17403; NID:G59591; PIDN:CA335325.1; PID:G1813973
;Note: possible protein-coding frames are given
;Note: the DNA sequence was submitted to EMBL, December 1989, in computer-readable form
;Genetics:
;Introns: 24/2; 86/1
;Superfamily: cytomagalovirus immediate-early protein
;Keywords: immediate-early protein

Query Match      100.0%; Score 41; DB 1; Length 491;
Best Local Similarity 100.0%; Pred. No. 0.61;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ELRRKMY 8
      |||||:|
Db      46 ELRRKLEY 53

RESULT 5
D69024
hypothetical protein MTH180 - Methanobacterium thermoautotrophicum (strain Delta H)
C.Species: Methanobacterium thermoautotrophicum
C.Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C.Accession: D69024
R.Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
Qiu, D.; Spadafora, R.; Vicair, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.
Ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A.Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: func
A.Reference number: A69000; MUID:98037514; PMID:9371463
A.Accession: D69024
A.Status: preliminary; nucleic acid sequence not shown; translation not shown
A.Molecule type: DNA
A.Residues: 1-305 <MTH>
A.Cross-references: GB:AB000886; GB:AE000666; NID:G2622276; PIDN:AA895669.1; PID:G262222;
A.Experimental source: strain Delta H
C.Genetics:
A.Gene: MTH180

Query Match      75.6%; Score 31; DB 2; Length 305;
Best Local Similarity 62.5%; Pred. No. 51;
Matches      5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 ELRRKMY 8
      |||||:|
Db      283 ELRRKMY 290

RESULT 6
T16406
hypothetical protein F48E3.7 - Caenorhabditis elegans
C.Species: Caenorhabditis elegans
C.Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Mar-2000
C.Accession: T16406
R.Fauley, A.
submitted to the EMBL Data Library, June 1995
A.Description: The sequence of C. elegans cosmid F48E3.
A.Reference number: Z18508
A.Accession: T16406
A.Status: preliminary; translated from GB/EMBL/DBJ
A.Molecule type: DNA
A.Residues: 1-426 <PAU>
A.Cross-references: EMBL:U28735; NID:G860708; PID:G860714; PIDN:AAA68268.1; CESP:F48E3.
A.Experimental source: strain Bristol N2
C.Genetics:
A.Gene: CESP:F48E3.7
A.Introns: 67/3; 119/2; 145/3; 192/3; 228/3; 269/3; 302/1; 354/3; 406/1
C.Superfamily: acetylcholine receptor

Query Match      75.6%; Score 31; DB 2; Length 426;
Best Local Similarity 75.0%; Pred. No. 72;
Matches      6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 ELRRKMY 8
      |||||:|
Db      229 ELRRKPLY 236

RESULT 7
G84866
hypothetical protein At2g34390 [imported] - Arabidopsis thaliana
C.Species: Arabidopsis thaliana (mouse-ear cress)
C.Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C.Accession: G84866
R.Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Unayam, L.; Tallon, L.

```



```

us, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
ature 402, 761-768, 1999
;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
;Reference number: A84420; MUID:20083487; PMID:10617197
;Accession: G84966
;Status: preliminary
;Molecule type: DNA
;Residues: 1-756 <STO>
;Cross-references: GB:AE002093; NID:92288988; PIDN:AAB64317.1; GSPDB:GN00139
;Genetics:
;Gene: At2g34390
;Map position: 2

Query Match 75.6%; Score 31; DB 2; Length 756;
Best Local Similarity 75.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

y 1 ELRRKMY 8
b 691 KLRRKMY 698

RESULT 8
9692
robable wall-associated kinase T4024.5 [imported] - Arabidopsis thaliana
;Species: Arabidopsis thaliana (mouse-ear cress)
;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
;Accession: E96692
;Theologian: A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
nson, N.F.; Hughes, B.; Huizar, L.
ature 408, 816-820, 2000
;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
;Li, J.H.; Li, X.; Liu, S.X.; Liu, Z.A.; Lurus, J.S.; Maiti, R.; Marziani,
izzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
er, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
;Reference number: A86141; MUID:21016719; PMID:11130712
;Accession: E96692
;Status: preliminary
;Molecule type: DNA
;Residues: 1-907 <STO>
;Cross-references: GB:AE005173; NID:gi1128388; PIDN:AAG31193.1; GSPDB:GN00141
;Genetics:
;Gene: T4024.5
;Map position: 1

Query Match 75.6%; Score 31; DB 2; Length 907;
Best Local Similarity 85.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

y 1 ELRRKMY 7
b 817 EVRRKMY 823

RESULT 9
ORTH
anthine dehydrogenase (EC 1.1.1.204) / xanthine oxidase {EC 1.1.3.22} - rat
;Alternate names: hypoxanthine oxidase
;Species: Rattus norvegicus (Norway rat)
;Date: 30-Apr-1991 #sequence_revision 07-Feb-1997 #text_change 19-Jan-2001
;Accession: A37810; S45259; S45260; S71397; 158308
;Anaya, Y.; Yamazaki, K.; Sato, M.; Noda, K.; Nishino, T.; Nishino, T.
;Biochem. 265, 14170-14175, 1990
;Title: Proteolytic conversion of xanthine dehydrogenase from the NAD-dependent type to
age sites of the enzyme protein during irreversible conversion by trypsin.
;Reference number: A37810; MUID:90354396; PMID:2387845
;Accession: A37810
;Molecule type: mRNA
;Residues: 1-478, 491-493, 'Q', 495-1331 <AMA>
;Cross-references: GB:J05579; NID:9207686

```

```

A;Note: parts of this sequence, including the amino end of the mature protein, were dete
R;Chow, C.W.; Clark, M.; Rinaldo, J.; Chalkley, R.
Nucleic Acids Res. 22, 1846-1854, 1994
;Title: Identification of the rat xanthine dehydrogenase/oxidase promoter.
;Reference number: 158308; MUID:94289906; PMID:8208609
;Accession: S45259
;Status: translated from GB/EMBL/DBJ
;Molecule type: mRNA
;Residues: 476-494 <RES>
;Cross-references: EMBL:U08123; NID:9473260; PIDN:AAB60444.1; PID:9473261
;Note: correction to A37810; sequence thought by authors to be macrophage splice form
;Accession: S45260
;Status: translation not shown
;Molecule type: DNA
;Residues: 1-55 <CHO>
;Cross-references: GB:U08122; NID:9472856; PIDN:AAA18869.1; PID:9472858; EMBL:U08121
R;Sato, A.; Nishino, T.; Noda, K.; Amaya, Y.; Nishino, T.
J. Biol. Chem. 270, 2818-2826, 1995
;Title: The structure of chicken liver xanthine dehydrogenase. cDNA cloning and the don
A;Reference number: A55711; MUID:9515354; PMID:7852355
A;Contents: annotation; confirmation of sequence
A;Note: the authors confirmed that both liver and macrophage mRNA's of the rat have a se
R;McManaman, J.L.; Shellman, V.; Wright, R.M.; Repine, J.E.
Arch. Biochem. Biophys. 332, 135-141, 1996
;Title: Purification of rat liver xanthine oxidase and xanthine dehydrogenase by affini
A;Reference number: S71397; MUID:96400342; PMID:8806718
;Accession: S71397
;Molecule type: protein
;Residues: 2-11 <MCM>
C;Comment: Xanthine dehydrogenase is reversibly converted to xanthine oxidase by oxidize
nversion to xanthine oxidase can also be performed artificially by a variety of sulfhydr
C;Comment: This enzyme contains four cofactors per subunit, one FAD, two iron-sulfur clu
C;Genetics:
A;Introns: 14/3; 34/1
A;Note: the list of introns may be incomplete
C;Complex: homodimer
C;Function: <XDH>
A;Description: catalyzes oxidation of xanthine to uric acid by NAD+ and water
A;Pathway: purine catabolism
C;Function: <XO>
A;Description: catalyzes oxidation of xanthine to uric acid and hydrogen peroxide by dic
C;Superfamily: xanthine dehydrogenase; ferredoxin [2Fe-2S] homology
C;Keywords: 2Fe-2S; FAD; flavoprotein; homodimer; iron-sulfur protein; metalloprotein; n
liem
F;2-1331/Product: xanthine dehydrogenase / xanthine oxidase #status experimental <AT>
F;26-74/Domain: ferredoxin [2Fe-2S] homology <FE1>
F;795-802/Region: nucleotide-binding motif A (P-loop)
F;43,48,51,73/Binding site: 2Fe-2S cluster (Cys) (covalent) #status predicted
F;112,115,147,149/Binding site: 2Fe-2S cluster (Cys) (covalent) #status predicted
F;825/Binding site: molybdopterin (Cys) (covalent) #status predicted
F;912/Binding site: molybdopterin (Arg) #status predicted
F;1261/Active site: Glu #status predicted

Query Match 75.6%; Score 31; DB 1; Length 1331;
Best Local Similarity 75.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ELRRKMY 8
DB 940 EVRRKMY 947

RESULT 10
XCMSDH
xanthine dehydrogenase (EC 1.1.1.204) / xanthine oxidase (EC 1.1.3.22) - mouse
;Alternate names: hypoxanthine oxidase
;Species: Mus musculus (house mouse)
;Date: 15-Mar-1996 #sequence_revision 07-Feb-1997 #text_change 19-Jan-2001
;Accession: 148374; S22419; S65134
R;Cazzaniga, G.; Terao, M.; Lo Schiavo, P.; Galbiati, F.; Segalla, P.; Seldin, M.F.; Gar
Genomics 23, 390-402, 1994
;Title: Chromosomal mapping, isolation, and characterization of the mouse xanthine dehy

```

```

;Reference number: A55561; MUID:95137585; PMID:7935888
;Accession: I48374
;Status: nucleic acid sequence not shown; translation not shown
;Molecule type: DNA
;Residues: 1-1335 <RES>
;Cross-references: EMBL:X75129; NID:9473040; PIDN:CAA52997.1; PID:g817959
;Note: the sequence and translation are shown only for the splice boundaries
;Terao, M.; Gazzaniga, G.; Ghezzi, P.; Bianchi, M.; Falciani, F.; Perani, P.; Garattini
Biochem. J. 283, 863-870, 1992
;Title: Molecular cloning of a cDNA coding for mouse liver xanthine dehydrogenase: regu
;Reference number: S22419; MUID:92272690; PMID:1590774
;Accession: S22419
;Molecule type: mRNA
;Residues: 1-240, '1', 242-620, 'M', 622-1335 <TER>
;Cross-references: EMBL:X62932; NID:g55443; PIDN:CAA44705.1; PID:g55444
;Ishii, T.; Aoki, N.; Noda, A.; Adachi, T.; Nakamura, R.; Matsuda, T.
Biochim. Biophys. Acta 1245, 285-292, 1995
;Title: Carboxy-terminal cytoplasmic domain of mouse butyrophilin specifically associat
;Reference number: S65133; MUID:96125722; PMID:8541302
;Accession: S65133
;Molecule type: Protein
;Residues: 2-9 <ISH>
;Comment: Xanthine dehydrogenase is reversibly converted to xanthine oxidase by oxidiz
;Version to xanthine oxidase can also be performed artificially by a variety of sulfhydry
;Comment: This enzyme contains four cofactors per subunit, one FAD, two iron-sulfur clu
;Genetics:
;Gene: XDH; XD; XO
;Introns: 17/3; 37/1; 69/2; 104/3; 147/1; 167/3; 190/3; 219/3; 267/1; 298/1; 348/3; 380
51/3; 1094/3; 1119/3; 1137/2; 1175/3; 1197/3; 1260/3; 1319/3
;Complex: homodimer
;Function: <XDH>
;Description: catalyzes oxidation of xanthine to uric acid by NAD+ and water
;Pathway: purine catabolism
;Function: <XO>
;Description: catalyzes oxidation of xanthine to uric acid and hydrogen peroxide by di
;Pathway: purine catabolism
;Superfamily: xanthine dehydrogenase; ferredoxin [2Fe-2S] homology
;Keywords: 2Fe-2S; FAD; flavoprotein; homodimer; iron-sulfur protein; metalloprotein; m
lism
?;2-1335/Product: xanthine dehydrogenase / xanthine oxidase #status predicted <MAT>
?;29-77/Domain: ferredoxin [2Fe-2S] homology <PER1>
?;798-805/Region: nucleotide-binding motif A (P-loop)
?;46-51.54.76/Binding site: 2Fe-2S cluster (Cys) (covalent) #status predicted
?;115-118.150.152/Binding site: 2Fe-2S cluster (Cys) (covalent) #status predicted
?;823/Binding site: molybdopterin (Cys) (covalent) #status predicted
?;915/Binding site: molybdopterin (Arg) #status predicted
?;1264/Active site: Glu #status predicted

Query Match 75.6%; Score 31; DB 1; Length 1335;
Best Local Similarity 75.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ELRRKMY 8
|:|:|:|
DB 943 EVRRKMY 950

RESULT 11
RVUNE
Genome polyprotein - Puumala virus (strain Hallnas 51)
N:Alternate names: L protein
N:Contains: RNA-directed RNA polymerase (BC 2.7.7.48)
C:Species: Puumala virus
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 29-May-1998
R:Stohwasser, R.; Raab, K.; Darai, G.; Bautz, E.K.F.
Virology 183, 386-391, 1991
;Title: Primary structure of the large (L) RNA segment of nephropathia epidemica virus
;Reference number: A40319; MUID:91272501; PMID:2053288
;Accession: A40319
;Molecule type: Genomic RNA
;Residues: 1-2156 <STO>
;Cross-references: GB:M63194

;Note: the source was designated as nephropathia epidemica virus
;Genetics:
;Map position: segment 1
;Superfamily: nephropathia epidemica virus RNA-directed RNA polymerase
;Keywords: nucleotidyltransferase

Query Match 75.6%; Score 31; DB 1; Length 2156;
Best Local Similarity 62.5%; Pred. No. 3.5e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ELRRKMY 8
|:|:|:|
DB 961 KLKRLMY 968

RESULT 12
D7221
transcription elongation factor, greB/greB family - Thermotoga maritima (strain MSB8)
C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: D7221
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hicke
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.
C.M.
Nature 399, 323-329, 1999
;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome se
;Reference number: A72200; MUID:99287316; PMID:10560571
;Accession: D7221
;Status: preliminary
;Molecule type: DNA
;Residues: 1-156 <ARN>
;Cross-references: GB:AE001810; GB:AE000512; NID:g4982271; PIDN:AA036773.1; PID:g49822
;Experimental source: strain MSB8
;Genetics:
;Gene: TM1706
;Superfamily: transcription elongation factor greB
;Keywords: transcription factor

Query Match 73.2%; Score 30; DB 2; Length 156;
Best Local Similarity 62.5%; Pred. No. 43;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ELRRKMY 8
|:|:|:|
DB 20 DLKGRFWY 27

RESULT 13
D81181
conserved hypothetical protein NMB0599 [imported] - Neisseria meningitidis (strain MC58)
C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C:Accession: D81181
R:Tetrelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.
xi, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; V
;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
;Reference number: A81000; MUID:20175755; PMID:10710307
;Accession: D81181
;Status: preliminary
;Molecule type: DNA
;Residues: 1-256 <RET>
;Cross-references: GB:AE002415; GB:AE002098; NID:g7225815; PIDN:AAF41027.1; PID:g72258.
;Experimental source: serogroup B, strain MC58
;Genetics:
;Gene: NMB0599
;Superfamily: conserved hypothetical protein H10188
C:Superfamily: conserved hypothetical protein H10188

Query Match 73.2%; Score 30; DB 2; Length 256;
Best Local Similarity 62.5%; Pred. No. 71;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

Y 1 ELRRKMMY 8
||||:|:
b 19 ELRRRLMW 26

RESULT 14

81925
robable sec-independent protein translocase component NMA0803 [imported] - Neisseria meningitidis
;Species: Neisseria meningitidis
;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
;Accession: C81925
;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel, Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream, et al. 404, 502-506, 2000
;Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
;Reference number: AB1775; MUID:20222536; PMID:10761919
;Accession: C81925
;Status: preliminary
;Molecule type: DNA
;Residues: 1-256 <PAR>
;Cross-references: GB:AL162754; GB:AL157959; NID:g7379424; PIDN:CAB84086.1; PID:g737952
;Experimental source: serogroup A, strain Z2491
;Genetics:
;Gene: tatC; NMA0803
;Superfamily: conserved hypothetical protein HI0188

Query Match 73.2%; Score 30; DB 2; Length 256;
Best Local Similarity 62.5%; Pred. No. 71;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Y 1 ELRRKMMY 8
||||:|:
b 19 ELRRRLMW 26

RESULT 15

I3081
conserved hypothetical protein Atu4277 [imported] - Agrobacterium tumefaciens (strain C58)
;Species: Agrobacterium tumefaciens
;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
;Accession: AI3081
;Wood, D.W.; Seubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.; Karp, P.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClelland, P.; Romero, P.; Zhang, S.
cience 294, 2317-2323, 2001
;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, ter, E.W.
;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
;Reference number: AB2577; MUID:21608550; PMID:11743193
;Accession: AI3081
;Status: preliminary
;Molecule type: DNA
;Residues: 1-263 <KUR>
;Cross-references: GB:AE008609; PIDN:AAL45071.1; PID:g1774273; CSPDB:GN00187
;Experimental source: strain C58 (Dupont)
;Genetics:
;Gene: Atu4277

Map position: linear chromosome
;Superfamily: Bacillus subtilis lactam utilization protein ycsF

Query Match 73.2%; Score 30; DB 2; Length 263;
Best Local Similarity 62.5%; Pred. No. 73;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Y 1 ELRRKMMY 8
||||:|:
b 90 ELRQOMLY 97

earch completed: May 21, 2004, 12:50:07
ob time : 10.3913 secs

GenCore version 5.1.6
 Copyright (c) 1993 - 2004 Compugen Ltd.
 W protein - protein search, using sw model
 run on: May 21, 2004, 12:34:48 ; Search time 7.30435 Seconds
 (without alignments)
 57.029 Million cell updates/sec
 title: US-09-980-058-9_COPY_1_8
 effect score: 41
 sequence: 1 ELRRKKMY 8
 coring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5
 earched: 141681 seqs, 52070155 residues
 total number of hits satisfying chosen parameters: 141681
 inimum DB seq length: 0
 aximum DB seq length: 2000000000
 ost-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries
 atabase : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

result No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	491	1	VIE1_HCMVA
2	41	100.0	491	1	VIE1_HCMVT
3	32	78.0	475	1	SYC_AERPE
4	31	75.6	102	1	RL3E_AERPE
5	31	75.6	430	1	TEK3_MOUSE
6	31	75.6	1330	1	XDH_RAT
7	31	75.6	1335	1	KDP_MOUSE
8	31	75.6	2156	1	RRPL_PUUMH
9	30	73.2	156	1	GREP_THEMA
10	30	73.2	181	1	RK5_GUTH
11	30	73.2	255	1	Y6H7_AGR7S
12	30	73.2	308	1	YEHX_SCOLI
13	30	73.2	371	1	RAL4_YEAST
14	30	73.2	459	1	K6PF_THELI
15	30	73.2	461	1	K6PF_THEZI
16	30	73.2	867	1	IF2_EUCAP
17	30	73.2	1073	1	PVD3_PLAKN
18	29	70.7	174	1	NADM_ARCFU
19	29	70.7	217	1	OGGI_FUSNN
20	29	70.7	380	1	OPCA_BACSU
21	29	70.7	381	1	OPBA_BACSU
22	29	70.7	440	1	GAAP_HUMAN
23	29	70.7	440	1	GAAP_RAT
24	29	70.7	457	1	SECY_CHLPN
25	29	70.7	473	1	MA1E_YEAST
26	29	70.7	489	1	EGAL_SULSH
27	29	70.7	556	1	SYFB_PYRHO
28	28	68.3	260	1	REDM_ECOLI
29	28	68.3	277	1	IF2A_AERPE
30	28	68.3	288	1	PNK_PORGI
31	28	68.3	368	1	ODPA_BACST
32	28	68.3	386	1	HISZ_RALSO
33	28	68.3	429	1	THIC_PYRAB

34 28 68.3 463 1 UHPT_ECOLI
 35 28 68.3 472 1 GATE_METJA
 36 28 68.3 473 1 MA6R_YEAST
 37 28 68.3 491 1 GATE_ANASP
 38 28 68.3 552 1 GAA4_MOUSE
 39 28 68.3 552 1 GAA4_RAT
 40 28 68.3 554 1 GAA4_HUMAN
 41 28 68.3 556 1 GAA4_BOVIN
 42 28 68.3 579 1 NH22_CAEEL
 43 28 68.3 759 1 MCM5_CAEEL
 44 28 68.3 763 1 PSRA_WOLSU
 45 28 68.3 803 1 Y920_TREPA

ALIGNMENTS

RESULT 1
 VIE1_HCMVA
 ID VIE1_HCMVA STANDARD; PRT; 491 AA.
 AC P13202;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 01-JUN-1993 (Rel. 26, Last annotation update)
 DE 55 kDa immediate-early protein 1 (IE1).
 GN U123.
 OS Human cytomegalovirus (strain AD169).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Cytomegalovirus.
 OX NCBI_TaxID=10360;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85194814; PubMed=2986374;
 RA Krigg A., Wilkinson G.W.G., Oram J.D.;
 RT "The structure of the major immediate early gene of human
 cytomegalovirus strain AD169.";
 RL Virus Res. 2:107-121(1985).
 RN [2]
 RP COMPLETE GENOME.
 RX MEDLINE=90269039; PubMed=2161319;
 RA Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,
 Horsnell T., Hutchison C.A. III, Kourazides T., Martignetti J.A.,
 Reddie E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;
 RT "Analysis of the protein-coding content of the sequence of human
 cytomegalovirus strain AD169.";
 RL Curr. Top. Microbiol. Immunol. 154:125-169(1990).
 CC -!- FUNCTION: THE ACTIVATION OF THE IE1.7 PROMOTER BY IE2 IS
 AUGMENTED BY THE IE1 PROTEIN.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X17403; CAA35325.1; --
 DR EMBL; M21293; AAA45980.1; --
 DR PIR; S09890; EDEEM5.
 KW Early protein; Phosphorylation; Nuclear protein.
 SQ SEQUENCE 491 AA; 55109 MW; CC5966B00CD8CB4 CRC64;

Query Match 100.0%; Score 41; DB 1; Length 491;
 Best Local Similarity 100.0%; Pred No. 0.22; Indels 0; Gaps 0;
 Matches 8; Conservative 0; Mismatches 0;
 Qy 1 ELRRKKMY 8
 |||||
 Db 199 ELRRKKMY 206

RESULT 2

```

VIE1_HCMVT
ID VIE1_HCMVT STANDARD; PRT; 491 AA.
AC P03169;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 55 kDa immediate-early protein 1 (IE1).
EN UL123.
DS Human cytomegalovirus (strain Towne).
DC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
DC Betaherpesvirinae; Cytomegalovirus.
DX NCBI_TaxID=10363;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84090395; PubMed=6317889;
3A Stenberg R.M., Thomsen D.R., Stinski M.F.;
RT "Structural analysis of the major immediate early gene of human
cytomegalovirus.";
RL J. Virol. 49:190-199(1984).
CC -!- FUNCTION: THE ACTIVATION OF THE IE1.7 PROMOTER BY IE1 IS
CC AUGMENTED BY THE IE1 PROTEIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M11630; AAA45979.1; -.
CC EXBL: M11628; AAA45979.1; JOINED.
CC EMBL: M11629; AAA45979.1; JOINED.
CC PIR: A03722; EDBIC.
KW Early protein; Phosphorylation; Nuclear protein.
FT CONFLICT 96 96 V -> L (IN REF. 1; AAA45979).
SQ SEQUENCE 491 AA; 55178 MW; E6041928A91A5867 CRC64;

Query Match 100.03; Score 41; DB 1; Length 491;
Best Local Similarity 100.03; Pred. No. 0.22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ELRRKQMY 8
DB 199 ELRRKQMY 206

RESULT 3
SVC_AERPE
ID SVC_AERPE STANDARD; PRT; 475 AA.
AC Q9YBK6;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE CysteinyI-trNA synthetase (EC 6.1.1.16) (Cysteine--trNA ligase)
DE (CYERS).
OS Aeropyrum pernix.
GN CYSS OR APE1592.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=99310339; PubMed=10382966;
RA Kwarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
crenarchaeon, Aeropyrum pernix K1.";
CC -!- SIMILARITY: Belongs to the L30E family of ribosomal proteins.

```

```

RL DNA Res. 6:83-101(1999).
CC -!- CATALYTIC ACTIVITY: ATP + L-cysteine + tRNA(Cys) = AMP +
CC diphosphate + L-cysteinyl-tRNA(Cys).
CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
CC Strong, to methionyl-tRNA synthetase.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AP000062; BAA80592.1; -.
CC PIR: C72538; C72538.
CC HAMAP: MP_00041; -.
CC InterPro: IPR002308; Cys_tRNA-synt_1a.
CC InterPro: IPR001412; tRNA-synt_1.
CC Pfam: PF01406; tRNA-synt_1e; 1.
CC PRINTS: PR00983; TRNASYNTHCS.
CC TIGRfams: TIGR00435; CYSS.1.
CC PROSITE: PS00178; AA_TRNA_LIGASE_1; FALSE NEG.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Metal-binding; Zinc; Complete proteome.
FT SITE 33 43 "HIGH" REGION.
FT SITE 267 271 "KMSKS" REGION.
FT BINDING 270 270 ATP (BY SIMILARITY).
FT METAL 31 31 ZINC (BY SIMILARITY).
FT METAL 210 210 ZINC (BY SIMILARITY).
FT METAL 235 235 ZINC (BY SIMILARITY).
FT METAL 239 239 ZINC (BY SIMILARITY).
SQ SEQUENCE 475 AA; 55746 MW; 99B8F7EA1698505D CRC64;

Query Match 78.03; Score 32; DB 1; Length 475;
Best Local Similarity 75.03; Pred. No. 19;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ELRRKQMY 8
DB 439 ELRRKQMY 446

RESULT 4
RL3E_AERPE
ID RL3E_AERPE STANDARD; PRT; 102 AA.
AC Q9YAU3;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L30e.
GN RPL30E OR APE1851.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=99310339; PubMed=10382966;
RA Kwarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
CC -!- SIMILARITY: Belongs to the L30E family of ribosomal proteins.

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; AF000062; BAA80855.1; -.
 CC PIR; B72571; B72571.
 CC HANAP; P14120; 1CN9.
 CC InterPro; IPR000231; Ribosomal_L30e.
 CC InterPro; IPR004038; Ribosomal_L7A.
 CC Pfam; PF01248; Ribosomal_L7Ae; 1.
 CC ProDom; PD004495; Ribosomal_L30e; 1.
 CC PROSITE; PS00709; Ribosomal_L30e; 1.
 CC PROSITE; PS00993; Ribosomal_L30e; 2; FALSE_NEG.
 CC Ribosomal protein; Complete proteome.
 CC SEQUENCE 102 AA; 11120 MW; 9349C5ED3BE3182 CRC64;

Query Match 75.6%; Score 31; DB 1; Length 102;
 Best Local Similarity 75.0%; Pred. No. 6.2; Indels 0; Gaps 0;
 Matches 6; Conservative 1; Mismatches 1;

iy 1 ELRRKMMY 3
 |||||:
 jb 46 ELRRKLEY 53

RESULT 5

ID TEK2 MOUSE STANDARD; PRT; 430 AA.
 CC Q922G7; Q9WVR0;
 CC 28-FEB-2003 (Rel. 41, Created)
 CC 28-FEB-2003 (Rel. 41, Last sequence update)
 CC 10-OCT-2003 (Rel. 42, Last annotation update)
 CC Tekin 2 (Tekin-t).
 CC TEK2.
 CC Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CC NCBI_TaxID=10090;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC TISSUE=Testis;
 CC MEDLINE=99383234; PubMed=10456331;
 CC Iguchi N., Tanaka H., Fujii T., Tamura K., Kaneko Y., Nojima H.,
 CC Nishimune Y.;
 CC "Molecular cloning of haploid germ cell-specific tekin cDNA and
 CC analysis of the protein in mouse testis.";
 CC FEBS Lett. 456:315-321(1999).
 CC [2]
 CC SEQUENCE FROM N.A.
 CC MEDLINE=22388257; PubMed=12477932;
 CC Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 CC Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
 CC Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 CC Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
 CC Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 CC Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
 CC Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
 CC Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 CC Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 CC Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 CC Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 CC Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
 CC Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 CC Blakesley R.W., Trichman J.W., Green E.D., Dickson M.C.,
 CC Rodriguez A.C., Grumman J.J., Schmutz J., Myers R.M.,
 CC Butterfield V.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 CC Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 CC "Generation and initial analysis of more than 15,000 full-length

RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -!- FUNCTION: Structural component of ciliary and flagellar
 CC microtubules. Forms filamentous polymers in the walls of ciliary
 CC and flagellar microtubules.
 CC -!- TISSUE SPECIFICITY: Testis specific.
 CC -!- DEVELOPMENTAL STAGE: Localized in the flagella of elongating
 CC spermatids from developmental step 15 to maturity.
 CC -!- SIMILARITY: Belongs to the tekin family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; AB027138; BAA81822.1; -.
 CC EMBL; BC008140; AAO8140.1; -.
 CC MGD; MGI:1346335; Tekt2.
 CC InterPro; IPR000435; Tektin.
 CC Pfam; PF03148; Tektin; 1.
 CC PRINIS; PR00511; TEKTIN.
 CC Microtubule; Structural protein; Coiled coil.
 CC DOMAIN 75 108 COILED COIL (POTENTIAL).
 CC DOMAIN 133 162 COILED COIL (POTENTIAL).
 CC DOMAIN 226 380 COILED COIL (POTENTIAL).
 CC CONFLICT 32 32 R -> Q (IN REF. 1).
 CC SEQUENCE 430 AA; 50311 MW; 1FE35ED77B5E9CDB CRC64;

Query Match 75.6%; Score 31; DB 1; Length 430;
 Best Local Similarity 85.7%; Pred. No. 28;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ELRRKMM 7
 |||||:
 Db 303 ELRRKMM 309

RESULT 6

ID XDH RAT STANDARD; PRT; 1330 AA.
 CC P22385; Q63157;
 CC 01-AUG-1991 (Rel. 19, Created)
 CC 01-FEB-1996 (Rel. 33, Last sequence update)
 CC 28-FEB-2003 (Rel. 41, Last annotation update)
 CC Xanthine dehydrogenase/oxidase [includes: Xanthine dehydrogenase
 CC (EC 1.1.1.204) (XD); Xanthine oxidase (EC 1.1.3.22) (XO) (Xanthine
 CC oxidoreductase)].
 CC XDH.
 CC Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 CC NCBI_TaxID=10116;
 CC [1]
 CC SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
 CC TISSUE=Liver;
 CC MEDLINE=90354396; PubMed=2387845;
 CC Amaya Y., Yamazaki K.-I., Sato M., Noda K., Nishino T., Nishino T.;
 CC "Proteolytic conversion of xanthine dehydrogenase from the
 CC NAD-dependent type to the O2-dependent type. Amino acid sequence of
 CC rat liver xanthine dehydrogenase and identification of the cleavage
 CC sites of the enzyme protein during irreversible conversion by
 CC trypsin.";
 CC J. Biol. Chem. 265:14170-14175(1990).
 CC [2]
 CC SEQUENCE OF 1-54 FROM N.A.
 CC STRAIN=Sprague-Dawley;
 CC MEDLINE=94269906; PubMed=8208609;
 CC Chow C.W., Clark M., Rinaldo J., Chalkley R.;
 CC "Identification of the rat xanthine dehydrogenase/oxidase promoter.";
 CC Nucleic Acids Res. 22:1846-1854(1994).

```

CC -!- FUNCTION: This enzyme can be converted from the dehydrogenase form
CC (D) to the oxidase form (O) irreversibly by proteolysis or
CC reversibly through the oxidation of sulphydryl groups.
CC -!- CATALYTIC ACTIVITY: Xanthine + NAD(+) + H(2)O = urate + NADH.
CC -!- CATALYTIC ACTIVITY: Xanthine + H(2)O + O(2) = urate + H(2)O(2).
CC -!- COFACTOR: FAD, molybdopterin, and two 2Fe-2S clusters.
CC -!- SUBUNIT: Homodimer.
CC -!- SUBCELLULAR LOCATION: Peroxisomal.
CC -!- INDUCTION: By interferon.
CC -!- SIMILARITY: TO 2FE-2S FERREDOXINS IN THE N-TERMINAL DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J05579; AAA42349.1; -.
CC EMBL; U08122; AAA18869.1; JOINED.
CC EMBL; U08120; AAA18869.1; JOINED.
CC EMBL; U08121; AAA18869.1; JOINED.
CC HSP; P80457; IPIQ.
CC InterPro; IPR002888; 2Fe-2S bind.
CC InterPro; IPR006058; 2Fe2S fd BS.
CC InterPro; IPR008274; Aldxan_dh_bind.
CC InterPro; IPR000674; Aldxan_dh_hamm.
CC InterPro; IPR005107; Co_dsh_flav_C.
CC InterPro; IPR002346; dehydrog_molyb.
CC InterPro; IPR001041; Ferredoxin.
CC InterPro; IPR000572; Oxidored_molyb.
CC Pfam; PF02738; Ald_Xan_dh_C2; 1.
CC Pfam; PF01315; Ald_Xan_dh_C; 1.
CC Pfam; PF03450; CO_dsh_Flav_C; 1.
CC Pfam; PF00941; FAD_binding_5; 1.
CC Pfam; PF00111; fer2; 1.
CC Pfam; PF01793; fer2; 2; 1.
CC ProDom; PD186071; 2Fe-2S bind; 1.
CC PROSITE; PS00197; 2FE2S FERREDOXIN; 1.
CC PROSITE; PS00559; MOLYBDOPTERIN_BUK; 1.
CC Oxidoreductase; NAD; Molybdenum; Flavoprotein; FAD; Metal-binding;
CC Iron-sulfur; Iron; 2Fe-2S.
CC INIT_MET 0
CC METAL 36 36 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC METAL 42 42 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC METAL 47 47 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC METAL 50 50 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC SEQUENCE 1330 AA; 146111 MW; A3DD206B9D74E565 CRC64;
CC -----
Query Match 75.68; Score 31; DB 1; Length 1330;
Best Local Similarity 75.08; Pred.No. 94;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
CC -----
CC 1 ELRRKMY 8
CC 939 EVRRKMY 946
CC -----
RESULT 7
CDH MOUSE
ID_KDH MOUSE STANDARD; PRT; 1335 AA.
AC Q00519;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Xanthine dehydrogenase/oxidase [includes: Xanthine dehydrogenase
DE (EC 1.1.1.204) (XD); Xanthine oxidase (EC 1.1.3.22) (XO) (Xanthine
DE oxidoreductase)].
EN KDH.
SS Mus musculus (Mouse).
SC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CC NCBI_faxID=10090;
CC (1)
CC SEQUENCE FROM N.A.
CC STRAIN=129/SV; TISSUE=Spleen;
CC MEDLINE=95137585; PubMed=7835888;
CC Cazzaniga G., Terao M., Lo Schiavo P., Galbiati F., Segalla F.,
CC Seldin M.P., Garattini E.;
CC "Chromosomal mapping, isolation, and characterization of the mouse
CC xanthine dehydrogenase gene.";
CC Genomics 23:390-402(1994).
CC (2)
CC SEQUENCE FROM N.A.
CC STRAIN=C57BL/6; TISSUE=Liver;
CC MEDLINE=9272690; PubMed=1590774;
CC Terao M., Cazzaniga G., Ghezzi P., Bianchi M., Falciani F.,
CC Perani P., Garattini E.;
CC "Molecular cloning of a cDNA coding for mouse liver xanthine
CC dehydrogenase. Regulation of its transcript by interferons in vivo.";
CC Biochem. J. 283:863-870(1992).
CC -!- FUNCTION: This enzyme can be converted from the dehydrogenase form
CC (D) to the oxidase form (O) irreversibly by proteolysis or
CC reversibly through the oxidation of sulphydryl groups.
CC -!- CATALYTIC ACTIVITY: Xanthine + NAD(+) + H(2)O = urate + NADH.
CC -!- CATALYTIC ACTIVITY: Xanthine + H(2)O + O(2) = urate + H(2)O(2).
CC -!- COFACTOR: FAD, molybdopterin, and two 2Fe-2S clusters.
CC -!- SUBUNIT: Homodimer.
CC -!- SUBCELLULAR LOCATION: Peroxisomal.
CC -!- INDUCTION: By interferon.
CC -!- SIMILARITY: Belongs to the xanthine dehydrogenase family.
CC -!- SIMILARITY: TO 2FE-2S FERREDOXINS IN THE N-TERMINAL DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X75129; CAA52997.1; JOINED.
CC EMBL; X75128; CAA52997.1; JOINED.
CC EMBL; X75127; CAA52997.1; JOINED.
CC EMBL; X75126; CAA52997.1; JOINED.
CC EMBL; X75125; CAA52997.1; JOINED.
CC EMBL; X75124; CAA52997.1; JOINED.
CC EMBL; X75123; CAA52997.1; JOINED.
CC EMBL; X75122; CAA52997.1; JOINED.
CC EMBL; X75121; CAA52997.1; JOINED.
CC EMBL; X75120; CAA52997.1; JOINED.
CC EMBL; X75119; CAA52997.1; JOINED.
CC EMBL; X75130; CAA52997.1; JOINED.
CC EMBL; X75131; CAA52997.1; JOINED.
CC EMBL; X75132; CAA52997.1; JOINED.
CC EMBL; X75133; CAA52997.1; JOINED.
CC EMBL; X75134; CAA52997.1; JOINED.
CC EMBL; X75135; CAA52997.1; JOINED.
CC EMBL; X75136; CAA52997.1; JOINED.
CC EMBL; X75137; CAA52997.1; JOINED.
CC EMBL; X75138; CAA52997.1; JOINED.
CC EMBL; X75139; CAA52997.1; JOINED.
CC EMBL; X75140; CAA52997.1; JOINED.
CC EMBL; X75141; CAA52997.1; JOINED.
CC EMBL; X75142; CAA52997.1; JOINED.
CC EMBL; X75143; CAA52997.1; JOINED.
CC EMBL; X75144; CAA52997.1; JOINED.
CC EMBL; X75145; CAA52997.1; JOINED.
CC EMBL; X75146; CAA52997.1; JOINED.
CC EMBL; X75147; CAA52997.1; JOINED.

```

```
R R EMBL; X75148; CAA52997.1; JOINED.
R R EMBL; X75149; CAA52997.1; JOINED.
R R EMBL; X75150; CAA52997.1; JOINED.
R R EMBL; X62932; CAA44705.1; -.
R R PIR; I48374; XOMSDH.
R R HSP; P80457; 1FO4.
R R MGD; MGI.98973; Xdh.
R R InterPro; IPR002888; 2Fe-2S bind.
R R InterPro; IPR006058; 2Fe2S_Fg_BS.
R R InterPro; IPR008274; Aldxan_dh_bind.
R R InterPro; IPR000674; Aldxan_dh_hamm.
R R InterPro; IPR005107; CO deh flav C.
R R InterPro; IPR002346; dehydrog molyb.
R R InterPro; IPR001041; Ferrredoxin.
R R InterPro; IPR000572; Oxidored molyb.
R R Pfam; PF02738; Ald_xan_dh_C2; 1.
R R Pfam; PF01315; Ald_xan_dh_C; 1.
R R Pfam; PF03450; CO deh flav C; 1.
R R Pfam; PF00941; FAD binding_5; 1.
R R Pfam; PF00111; fer2; 1.
R R Pfam; PF01799; fer2_2; 1.
R R ProDom; PD186071; 2Fe-2S bind; 1.
R R PROSITE; PS00197; 2FE2S_FERRDOXIN; 1.
R R PROSITE; PS00559; MOLYDOPTERIN_EUK; 1.
R R Oxidoreductase; NAD; Molybdenum; Flavoprotein; FAD; Metal-binding;
R R Iron-sulfur; Iron; 2Fe-2S.
T T METAL 40 40 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
T T METAL 46 46 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
T T METAL 51 51 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
T T METAL 54 54 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
T T CONFLICT 241 241 V -> I (IN REF. 2).
T T CONFLICT 621 621 T -> M (IN REF. 2).
T T SEQUENCE 1335 AA; 146517 MW; 99CE6FD8B42F858 CRC64;

Query Match 75.68; Score 31; DB 1; Length 1335;
Best Local Similarity 75.08; Pred. NO. 94;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Y 1 ELRRKMY 8
b 943 EVRRKMY 950

RESULT 8
RPL_PUUMH STANDARD; PRT; 2156 AA.
CC P27176;
CC 01-AUG-1992 (Rel. 23, Created)
CC 01-AUG-1992 (Rel. 23, Last sequence update)
CC 01-AUG-1992 (Rel. 23, Last annotation update)
CC RNA-directed RNA polymerase (EC 2.7.7.48) (L protein).
CC L.
CC Puumala virus (strain Hallnas B1) (Nephropathia epidemica virus).
CC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Hantavirus.
CC NCBI_TaxID=11605;
CC (1)
CC SEQUENCE FROM N.A.
CC MEDLINE=9127501; PubMed=2053288;
CC Steinhasser R., Raab K., Darai G., Bantz E.K.P.;
CC "Primary structure of the large (L) RNA segment of nephropathia
CC epidemica virus strain Hallnas B1 coding for the viral RNA
CC polymerase."
CC Virology 183:386-391(1991).
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate +
CC [RNA] (N).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X75148; CAA52997.1; JOINED.
CC PIR; I48374; XOMSDH.
CC HSP; P80457; 1FO4.
CC MGD; MGI.98973; Xdh.
CC InterPro; IPR002888; 2Fe-2S bind.
CC InterPro; IPR006058; 2Fe2S_Fg_BS.
CC InterPro; IPR008274; Aldxan_dh_bind.
CC InterPro; IPR000674; Aldxan_dh_hamm.
CC InterPro; IPR005107; CO deh flav C.
CC InterPro; IPR002346; dehydrog molyb.
CC InterPro; IPR001041; Ferrredoxin.
CC InterPro; IPR000572; Oxidored molyb.
CC Pfam; PF02738; Ald_xan_dh_C2; 1.
CC Pfam; PF01315; Ald_xan_dh_C; 1.
CC Pfam; PF03450; CO deh flav C; 1.
CC Pfam; PF00941; FAD binding_5; 1.
CC Pfam; PF00111; fer2; 1.
CC Pfam; PF01799; fer2_2; 1.
CC ProDom; PD186071; 2Fe-2S bind; 1.
CC PROSITE; PS00197; 2FE2S_FERRDOXIN; 1.
CC PROSITE; PS00559; MOLYDOPTERIN_EUK; 1.
CC Oxidoreductase; NAD; Molybdenum; Flavoprotein; FAD; Metal-binding;
CC Iron-sulfur; Iron; 2Fe-2S.
CC METAL 40 40 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC METAL 46 46 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC METAL 51 51 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC METAL 54 54 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC CONFLICT 241 241 V -> I (IN REF. 2).
CC CONFLICT 621 621 T -> M (IN REF. 2).
CC SEQUENCE 1335 AA; 146517 MW; 99CE6FD8B42F858 CRC64;

Query Match 75.68; Score 31; DB 1; Length 1335;
Best Local Similarity 75.08; Pred. NO. 94;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Y 1 ELRRKMY 8
b 943 EVRRKMY 950

RESULT 9
GREA_THEME STANDARD; PRT; 156 AA.
CC Q9X232;
CC 30-MAY-2000 (Rel. 39, Created)
CC 30-MAY-2000 (Rel. 39, Last sequence update)
CC 15-MAR-2004 (Rel. 43, Last annotation update)
CC DE Transcription elongation factor greA (Transcript cleavage factor
CC greA).
CC GN GREA OR TM1706.
CC OS Thermotoga maritima.
CC OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
CC OX NCBI_TaxID=23356;
CC (1)
CC SEQUENCE FROM N.A.
CC STRAIN=MSB8 / DSM 3109 / ATCC 43589;
CC MEDLINE=99287316; PubMed=10360571;
CC Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
CC Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
CC McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
CC Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
CC Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
CC Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
CC "Evidence for lateral gene transfer between Archaea and Bacteria from
CC genome sequence of Thermotoga maritima."
CC Nature 398:323-329(1999).
CC -!- FUNCTION: Necessary for efficient RNA polymerase transcription
CC elongation past template-encoded arresting sites. The arresting
CC sites in DNA have the property of trapping a certain fraction of
CC elongating RNA polymerases that pass through, resulting in locked
CC ternary complexes. Cleavage of the nascent transcript by cleavage
CC factors such as greA or greB allows the resumption of elongation
CC from the new 3'terminus. GreA releases sequences of 2 to 3
CC nucleotides (By similarity).
CC -!- SIMILARITY: Belongs to the greA/greB family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; A5001810; AAD36773.1; -.
CC PIR; D72221; D72221.
CC HSSP; P21346; 1GRJ.
CC TIGR; TM1706; -.
CC HANAP; MF_00105; -.
CC InterPro; IPR006359; GreA.
CC InterPro; IPR001437; GreA_GreB.
CC Pfam; PF01272; GreA_GreB; 1.
CC Pfam; PF03449; GreA_GreB_N; 1.
CC ProDom; PD004918; GreA_GreB; 1.
CC TIGRFAMs; TIGR01462; GreA; 1.
```



```

EHX_ECOLI
D YEHX_ECOLI STANDARD; PRT; 308 AA.
C P33360;
T 01-FEB-1994 (Rel. 28, Created)
T 01-FEB-1994 (Rel. 28, Last sequence update)
T 16-OCT-2001 (Rel. 40, Last annotation update)
E Hypothetical ABC transporter ATP-binding protein yehX.
N YEHX OR B2129
S Escherichia coli.
C Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
C Enterobacteriaceae; Escherichia.
X NCBI_TaxID=562;
N [1]
P SEQUENCE FROM N.A.
C STRAIN=K12 / BHE2600;
A Richerich P., Lakey N., Gryan G., Jaehn L., Mintz L., Robison K.,
A Church G.M.;
L Submitted (OCT-1993) to the EMBL/GenBank/DDBJ databases.
N [2]
P SEQUENCE FROM N.A.
C STRAIN=K12 / MG1655;
X MEDLINE=97426617; PubMed=9278503;
A Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
A Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
A Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
A Mau B., Shao Y.;
T "The complete genome sequence of Escherichia coli K-12.";
L Science 277:1453-1474 (1997).
C -!- SIMILARITY: Belongs to the ABC transporter family.
-----
C This SWISS-PROT entry is copyright. It is produced through a collaboration
C between the Swiss Institute of Bioinformatics and the EMBL outstation -
C the European Bioinformatics Institute. There are no restrictions on its
C use by non-profit institutions as long as its content is in no way
C modified and this statement is not removed. Usage by and for commercial
C entities requires a license agreement (See http://www.isb-sib.ch/announce/
C or send an email to license@isb-sib.ch).
-----
C EMBL; U00007; AAA60492.1; -.
C EMBL; A300302; AAC75190.1; -.
R PIR; H64980; H64980.
R EcoGene; EG12010; YehX.
R InterPro; IPR003593; AAA ATPase.
R InterPro; IPR003439; ABC transporter.
R InterPro; IPR000644; CBS domain.
R Pfam; PF00005; ABC_tran; 1.
R ProDom; PD000006; ABC transporter; 1.
R SMART; SM00382; AAA; 1.
R PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
R PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
W Hypothetical protein; ATP-binding; Transport; Complete proteome.
T NP BIND 34 41 ATP (BY SIMILARITY).
Q SEQUENCE 308 AA; 34424 MW; 037C7C94E4E6FB690 CRC64;

Query Match 73.2%; Score 30; DB 1; Length 308;
Best Local Similarity 75.0%; Pred. No. 33;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Y 1 ELRRKMY 8
b 72 ELRRRMY 79
|||||
72 ELRRRMY 79

RESULT 13
A14 YEAST STANDARD; PRT; 371 AA.
D RAI14 YEAST
C P28519;
T 01-DEC-1992 (Rel. 24, Created)
T 01-FEB-1996 (Rel. 33, Last sequence update)
T 10-OCT-2003 (Rel. 42, Last annotation update)
E DNA repair protein RAD14.
N RAD14 OR YMR201C OR YMR325.02C.

```

```

OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
N [1]
P SEQUENCE FROM N.A., AND MUTANT RAD14-2.
X MEDLINE=97197968; PubMed=9046084;
A Jones G.W., Reed S.H., Waters R.;
R "Characterization of the rad14-2 mutant of Saccharomyces cerevisiae:
R implications for the recognition of UV photoproducts by the Rad14
R protein.";
RL Yeast 13:31-36 (1997).
N [2]
P SEQUENCE OF 125-371 FROM N.A.
X MEDLINE=92158051; PubMed=1741034;
A Bankmann M., Prakash L., Prakash S.;
R "Yeast RAD14 and human Xeroderma pigmentosum group A DNA-repair genes
R encode homologous proteins.";
RL Nature 355:555-558 (1992).
N [3]
P SEQUENCE FROM N.A.
X STRAIN=S288C / AB972;
R MEDLINE=97333268; PubMed=9169872;
A Bowman S., Churchman C.M., Badcock K., Brown D., Chillingworth T.,
A Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
A Jagels K., Iye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
A Rice P., Skelton J., Walsh S., Whitehead S., Barrell B.G.;
R "The nucleotide sequence of Saccharomyces cerevisiae chromosome
R XIII.";
RL Nature 387:90-93 (1997).
N [4]
P CHARACTERIZATION.
X MEDLINE=93296152; PubMed=8516285;
R Guider S.N., Sung P., Prakash L., Prakash S.;
R "Yeast DNA-repair gene RAD14 encodes a zinc metalloprotein with
R affinity for ultraviolet-damaged DNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:5433-5437 (1993).
CC -!- FUNCTION: INVOLVED IN NUCLEOTIDE EXCISION REPAIR. BINDS
CC SPECIFICALLY TO DAMAGED DNA. REQUIRED FOR THE INCISION STEP.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Belongs to the XPA family.
-----
C This SWISS-PROT entry is copyright. It is produced through a collaboration
C between the Swiss Institute of Bioinformatics and the EMBL outstation -
C the European Bioinformatics Institute. There are no restrictions on its
C use by non-profit institutions as long as its content is in no way
C modified and this statement is not removed. Usage by and for commercial
C entities requires a license agreement (See http://www.isb-sib.ch/announce/
C or send an email to license@isb-sib.ch).
-----
C EMBL; X64064; CAA5420.1; -.
C EMBL; Z48755; CAA88642.1; -.
R PIR; S59442; S59442.
R GenOnline; 142876; -.
R SGD; S0004814; RAD14.
R InterPro; IPR000465; XPA_protein.
R Pfam; PF05181; XPA_C; 1.
R TIGRfams; TIGR00598; rad14; 1.
R PROSITE; PS00752; XPA_1; 1.
R PROSITE; PS00753; XPA_2; 1.
KX DNA repair; DNA-binding; Zinc-finger; Nuclear protein.
FT ZN FING 191 216
FT MUTAGEN 207 207
FT MUTAGEN 216 216
V->M: IN RAD14-2; LOSS OF RECOGNITION OF
CYCLOBUTANE PYRIMIDINE DIMERS.
C->Y: IN RAD14-2; LOSS OF RECOGNITION OF
CYCLOBUTANE PYRIMIDINE DIMERS.
FT MUTAGEN 216 216
SQ SEQUENCE 371 AA; 43038 MW; A90FA9CAAGF3EA52 CRC64;

Query Match 73.2%; Score 30; DB 1; Length 371;
Best Local Similarity 62.5%; Pred. No. 40;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ELRRKMY 8

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

CM protein - protein search, using sw model

Run on: May 21, 2004, 12:43:34 ; Search time 25.3913 Seconds
(without alignments)
99.410 Million cell updates/sec

Title: US-09-980-058-9_COPY_1_8
Perfect score: 41
Sequence: 1 ELRKKMY 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREXEL 25:.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp Vertebrate.*
14: sp Unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	99	12 Q9YRK5	Q9YRK5 human cytom
2	41	100.0	100	12 Q9W9C4	Q9W9C4 human cytom
3	41	100.0	100	12 Q9YRL5	Q9YRL5 human cytom
4	41	100.0	100	12 Q9YRL1	Q9YRL1 human cytom
5	41	100.0	100	12 Q9YRL0	Q9YRL0 human cytom
6	41	100.0	101	12 Q9WKR9	Q9WKR9 human cytom
7	41	100.0	104	12 Q9WKS1	Q9WKS1 human cytom
8	41	100.0	105	12 Q9WKS7	Q9WKS7 human cytom
9	41	100.0	105	12 Q9YRL9	Q9YRL9 human cytom
10	41	100.0	107	12 Q9YRL0	Q9YRL0 human cytom
11	41	100.0	107	12 Q9WKS2	Q9WKS2 human cytom
12	41	100.0	108	12 Q9YRK7	Q9YRK7 human cytom
13	41	100.0	110	12 Q9YRK9	Q9YRK9 human cytom
14	41	100.0	111	12 Q9W8Q5	Q9W8Q5 human cytom
15	41	100.0	111	12 Q9YRK6	Q9YRK6 human cytom
16	38	92.7	100	12 Q9YRL4	Q9YRL4 human cytom

17	38	92.7	109	12	Q9YRM2	Q9YRM2 human cytom
18	38	92.7	110	12	Q9YRM1	Q9YRM1 human cytom
19	38	92.7	112	12	Q9YRL2	Q9YRL2 human cytom
20	38	92.7	125	12	Q9YRK3	Q9YRK3 human cytom
21	37	90.2	103	12	Q9YRK8	Q9YRK8 human cytom
22	37	90.2	105	12	Q9YRL7	Q9YRL7 human cytom
23	34	82.9	100	12	Q9W8G6	Q9W8G6 human cytom
24	34	82.9	103	12	Q9WKS4	Q9WKS4 human cytom
25	34	82.9	105	12	Q9WKS6	Q9WKS6 human cytom
26	34	82.9	105	12	Q9YRK2	Q9YRK2 human cytom
27	34	82.9	105	12	Q9YRK1	Q9YRK1 human cytom
28	34	82.9	108	12	Q9WKS5	Q9WKS5 human cytom
29	34	82.9	108	12	Q9WKS0	Q9WKS0 human cytom
30	34	82.9	110	12	Q9YRL3	Q9YRL3 human cytom
31	34	82.9	110	12	Q9W9A8	Q9W9A8 human cytom
32	34	82.9	111	12	Q9YJR1	Q9YJR1 human cytom
33	33	80.5	109	5	Q9V389	Q9V389 encephalito
34	33	80.5	363	5	Q8SR94	Q8SR94 drosophila
35	33	80.5	426	5	Q8IR17	Q8IR17 drosophila
36	33	80.5	568	16	Q8EKL7	Q8EKL7 shewanella
37	33	80.5	746	3	Q9HG15	Q9HG15 collettotic
38	32	78.0	100	12	Q9WKS3	Q9WKS3 human cytom
39	32	78.0	106	12	Q9YRK4	Q9YRK4 human cytom
40	32	78.0	106	12	Q9YRK0	Q9YRK0 human cytom
41	32	78.0	111	12	Q9YRL6	Q9YRL6 human cytom
42	32	78.0	111	12	Q9YRL8	Q9YRL8 human cytom
43	32	78.0	338	13	Q7SZ60	Q7SZ60 brachydanio
44	32	78.0	362	4	Q96TAS	Q96TAS homo sapien
45	32	78.0	404	12	Q911W4	Q911W4 pea seed-bo

ALIGNMENTS

RESULT 1

ID Q9YRK5 PRELIMINARY; PRT; 99 AA.
AC Q9YRK5;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Major immediate-early protein (Fragment).
OS Human cytomegalovirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10359;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1343/93;
RX MEDLINE=99036778; PubMed=8817892;
RA Zwegberg Wirgatt B., Brytting M., Linde A., Wahren B., Grillner L.;
RT "Sequence variation within three important cytomegalovirus gene
regions in isolates from four different patient populations.";
RL J. Clin. Microbiol. 36:3662-3669(1998).
DR EMBL, AF099584; AAD04506.1; --
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002198; ADH short.
DR PROSITE; PS00061; ADH SHORT; 1.
FT NON_TER 1
FT NON_TER 99
SQ SEQUENCE 99 AA; 11420 MW; B425PEP4D53F28B4 CRC64;

Query Match 100.0%; Score 41; DB 12; Length 99;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ELRKKMY 8
Db 50 ELRKKMY 57

RESULT 2

RESULT	4
ID	Q9YRL1
NAME	PRELIMINARY; PRT; 100 AA.
AC	Q9YRL1
DT	01-MAY-1999 (TrEMBLrel. 10, Created)
DT	01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE	Major immediate-early protein (Fragment).
DE	Human cytomegalovirus.
OS	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC	Betaherpesvirinae; Cytomegalovirus.
OX	NCBI_TaxID=10359;
RN	[1]
RP	SEQUENCE FROM N.A.
KP	STRAIN=589/91;
RC	MEDLINE=99036778; PubMed=9817892;
RA	Zweygberg Wartat B., Brytting M., Linde A., Wahren B., Grillner L.;
RA	"Sequence variation within three important cytomegalovirus gene
RT	"regions in isolates from four different patient populations.";
RT	J. Clin. Microbiol. 36:3662-3669(1998).
DR	EMB1; AF099577; AAD04499.1; -
FT	NON TER 1
FT	NON TER 100
SQ	SEQUENCE 100 AA; 11402 MW; 77AB4A9C95878C00 CRC64;
Query Match	100.0%; Score 41; DB 12; Length 100;
Best Local Similarity	100.0%; Pred. No. 0.28; Indels 0; Gaps 0;
Matches	8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 ELRRKMY 8
DB	51 ELRRKMY 58

```

RESULTS
Q9YRLO PRELIMINARY; PRT; 100 AA.
ID Q9YRLO;
AC Q9YRLO;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Major immediate-early protein [Fragment].
OS Human cytomegalovirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
NCBI TaxID=10359;
f1;_taxid=10359;
RN f1;
RP SEQUENCE FROM N.A.
RC STRAIN=45/92;
RX MEDLINE=99036778; PubMed=9817892;
RA zweyberg Wargat B., Brytting M., Linde A., Wahren B., Grillner L.;
RT "Sequence variation within three important cytomegalovirus gene
RL regions in isolates from four different patient populations.";
RT J. Clin. Microbiol. 36:3662-3669(1998).
DR EMBL; AF093579; AA004501.1; -.
DR GO; GO:0016491; P:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002198; ADH short.
DR PROSITE; PS00063; ADH SHORT; 1.
PT NON TER 1
FT NON TER 100
SQ SEQUENCE 100 AA; 11471 MW; 7862349CE2FID684 CRC64;

Query Match 100.0%; Score 41; DB 12; Length 100;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 BLKKNKY 8
| | | | |
db 51 ELKKKKY 58

```

9WKS9
D Q9WKS9 PRELIMINARY; PRT; 101 AA.
C Q9WKS9
T 01-NOV-1999 (TREMELrel. 12, Created)
T 01-NOV-1999 (TREMELrel. 12, Last sequence update)
T 01-JUN-2003 (TREMELrel. 24, Last annotation update)
E Major immediate-early protein (Fragment).
S Human cytomegalovirus.
C Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
C Betaherpesvirinae; Cytomegalovirus.
X NCBI_TaxID=10359;
N [1]
P SEQUENCE FROM N.A.
C STRAIN=377/94;
X MEDLINE=99036778; PubMed=9817892;
A Zeyberg Wirgart B., Brytting M., Linde A., Wahren B., Grillner L.;
T "Sequence variation within three important cytomegalovirus gene
regions in isolates from four different patient populations.";
L J. Clin. Microbiol. 36:3662-3669(1998).
R EMBL: AF099596; AAD04518.1; -.
R GO: GO:0016491; F:oxidoreductase activity; IEA.
R GO: GO:0008152; P:metabolism; IEA.
R InterPro: IPR002198; ADH_SHORT; 1.
R PROSITE: PS00061; ADH_SHORT; 1.
T NON TER 1
T NON TER 101
Q SEQUENCE 101 AA; 11653 MW; A588677188B6FID6 CRC64;

Query Match 100.0%; Score 41; DB 12; Length 101;
Best Local Similarity 100.0%; Pred. No. 0.28; Mismatches 0; Indels 0; Gaps 0;
Matches 8; Conservative 0

Y 1 ELRRKMY 8
b 51 ELRRKMY 58

RESULT 7
9WKS1
D Q9WKS1 PRELIMINARY; PRT; 104 AA.
C Q9WKS1
T 01-NOV-1999 (TREMELrel. 12, Created)
T 01-NOV-1999 (TREMELrel. 12, Last sequence update)
T 01-DEC-2001 (TREMELrel. 19, Last annotation update)
E Major immediate-early protein (Fragment).
S Human cytomegalovirus.
C Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
C Betaherpesvirinae; Cytomegalovirus.
X NCBI_TaxID=10359;
N [1]
P SEQUENCE FROM N.A.
C STRAIN=454/94;
X MEDLINE=99036778; PubMed=9817892;
A Zeyberg Wirgart B., Brytting M., Linde A., Wahren B., Grillner L.;
T "Sequence variation within three important cytomegalovirus gene
regions in isolates from four different patient populations.";
L J. Clin. Microbiol. 36:3662-3669(1998).
R EMBL: AF099587; AAD04509.1; -.
R GO: GO:0016491; F:oxidoreductase activity; IEA.
R GO: GO:0008152; P:metabolism; IEA.
R InterPro: IPR002198; ADH_SHORT; 1.
R PROSITE: PS00061; ADH_SHORT; 1.
T NON TER 1
T NON TER 104
Q SEQUENCE 104 AA; 11911 MW; D64A1E8B47AE0F88 CRC64;

Query Match 100.0%; Score 41; DB 12; Length 104;
Best Local Similarity 100.0%; Pred. No. 0.29; Mismatches 0; Indels 0; Gaps 0;
Matches 8; Conservative 0

Y 1 ELRRKMY 8
b 51 ELRRKMY 58

RESULT 8
9WKS7

9WKS7
D Q9WKS7 PRELIMINARY; PRT; 105 AA.
C Q9WKS7
T 01-NOV-1999 (TREMELrel. 12, Created)
T 01-NOV-1999 (TREMELrel. 12, Last sequence update)
T 01-DEC-2001 (TREMELrel. 19, Last annotation update)
E Major immediate-early protein (Fragment).
S Human cytomegalovirus.
C Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
C Betaherpesvirinae; Cytomegalovirus.
X NCBI_TaxID=10359;
N [1]
P SEQUENCE FROM N.A.
C STRAIN=1606/84;
X MEDLINE=99036778; PubMed=9817892;
A Zeyberg Wirgart B., Brytting M., Linde A., Wahren B., Grillner L.;
T "Sequence variation within three important cytomegalovirus gene
regions in isolates from four different patient populations.";
L J. Clin. Microbiol. 36:3662-3669(1998).
R EMBL: AF099558; AAD04480.1; -.
R GO: GO:0016491; F:oxidoreductase activity; IEA.
R GO: GO:0008152; P:metabolism; IEA.
R InterPro: IPR002198; ADH_SHORT; 1.
R PROSITE: PS00061; ADH_SHORT; 1.
T NON TER 1
T NON TER 105
Q SEQUENCE 105 AA; 12024 MW; B5664A1E8B47AE0F CRC64;

Query Match 100.0%; Score 41; DB 12; Length 105;
Best Local Similarity 100.0%; Pred. No. 0.29; Mismatches 0; Indels 0; Gaps 0;
Matches 8; Conservative 0

Y 1 ELRRKMY 8
b 51 ELRRKMY 58

RESULT 9
9WKS9
D Q9WKS9 PRELIMINARY; PRT; 105 AA.
C Q9WKS9
T 01-NOV-1999 (TREMELrel. 10, Created)
T 01-MAY-1999 (TREMELrel. 10, Last sequence update)
T 01-JUN-2003 (TREMELrel. 24, Last annotation update)
E Major immediate-early protein (Fragment).
S Human cytomegalovirus.
C Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
C Betaherpesvirinae; Cytomegalovirus.
X NCBI_TaxID=10359;
N [1]
P SEQUENCE FROM N.A.
C STRAIN=3164/93;
X MEDLINE=99036778; PubMed=9817892;
A Zeyberg Wirgart B., Brytting M., Linde A., Wahren B., Grillner L.;
T "Sequence variation within three important cytomegalovirus gene
regions in isolates from four different patient populations.";
L J. Clin. Microbiol. 36:3662-3669(1998).
R EMBL: AF099562; AAD04484.1; -.
R GO: GO:0016491; F:oxidoreductase activity; IEA.
R GO: GO:0008152; P:metabolism; IEA.
R InterPro: IPR002198; ADH_SHORT; 1.
R PROSITE: PS00061; ADH_SHORT; 1.
T NON TER 1
T NON TER 105
Q SEQUENCE 105 AA; 12063 MW; B5231EC7B5941771 CRC64;

Query Match 100.0%; Score 41; DB 12; Length 105;
Best Local Similarity 100.0%; Pred. No. 0.29; Mismatches 0; Indels 0; Gaps 0;
Matches 8; Conservative 0

Y 1 ELRRKMY 8
b 51 ELRRKMY 58

RESULT 10
9WKS0
D Q9WKS0 PRELIMINARY; PRT; 107 AA.

```

YC Q9YRMO;
YT 01-MAY-1999 (TrEMBLrel. 10, Created)
YT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
YT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Major immediate-early protein (Fragment).
DE Human cytomegalovirus.
XC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
XX NCBI_TaxID=10359;
XX [1]
XX SEQUENCE FROM N.A.
YC STRAIN=1577/84;
YC MEDLINE=99036778; PubMed=9817892;
YA Zwegberg Wirgart B., Brytting M., Linde A., Wahren B., Grillner L.;
YT "Sequence variation within three important cytomegalovirus gene
YT regions in isolates from four different patient populations.";
YL J. Clin. Microbiol. 36:3662-3669(1998).
XR EMBL; AF099557; AAD04479.1; -.
YT NON_TER 1
YT NON_TER 107
YT SEQUENCE 107 AA; 12315 MW; DF1085664A1E8B5F CRC64;

Query Match 100.0%; Score 41; DB 12; Length 107;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 ELRRKMY 8
Y 51 ELRRKMY 58

RESULT 11
Q9YRMO;
ID Q9YRMO; PRELIMINARY; PRT; 107 AA.
AC Q9YRMO;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Major immediate-early protein (Fragment).
DE Human cytomegalovirus.
XC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
XX NCBI_TaxID=10359;
XX [1]
XX SEQUENCE FROM N.A.
YC STRAIN=829/91;
YC MEDLINE=99036778; PubMed=9817892;
YA Zwegberg Wirgart B., Brytting M., Linde A., Wahren B., Grillner L.;
YT "Sequence variation within three important cytomegalovirus gene
YT regions in isolates from four different patient populations.";
YL J. Clin. Microbiol. 36:3662-3669(1998).
XR EMBL; AF099578; AAD04500.1; -.
YT NON_TER 1
YT NON_TER 107
YT SEQUENCE 107 AA; 12299 MW; 851085664A1E8B47 CRC64;

Query Match 100.0%; Score 41; DB 12; Length 107;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 ELRRKMY 8
Y 51 ELRRKMY 58

RESULT 12
Q9YRMO;
ID Q9YRMO; PRELIMINARY; PRT; 108 AA.
AC Q9YRMO;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Major immediate-early protein (Fragment).
DE Human cytomegalovirus.

```

```

OS Human cytomegalovirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10359;
OX [1]
OX SEQUENCE FROM N.A.
RP STRAIN=61/92;
RX MEDLINE=99036778; PubMed=9817892;
RA Zwegberg Wirgart B., Brytting M., Linde A., Wahren B., Grillner L.;
RT "Sequence variation within three important cytomegalovirus gene
RT regions in isolates from four different patient populations.";
RL J. Clin. Microbiol. 36:3662-3669(1998).
DR EMBL; AF099582; AAD04504.1; -.
YT NON_TER 1
YT NON_TER 108
YT SEQUENCE 108 AA; 12498 MW; 3F371862997FDEC9 CRC64;

Query Match 100.0%; Score 41; DB 12; Length 108;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 ELRRKMY 8
Y 51 ELRRKMY 58

RESULT 13
Q9YRMO;
ID Q9YRMO; PRELIMINARY; PRT; 110 AA.
AC Q9YRMO;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Major immediate-early protein (Fragment).
DE Human cytomegalovirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10359;
OX [1]
OX SEQUENCE FROM N.A.
YC STRAIN=410/94;
YC MEDLINE=99036778; PubMed=9817892;
YA Zwegberg Wirgart B., Brytting M., Linde A., Wahren B., Grillner L.;
RT "Sequence variation within three important cytomegalovirus gene
RT regions in isolates from four different patient populations.";
RL J. Clin. Microbiol. 36:3662-3669(1998).
DR EMBL; AF099597; AAD04519.1; -.
DR GO; GO:0016491; P:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002198; ADH_short.
DR PROSITE; PS00061; ADH_SHORT; 1.
YT NON_TER 1
YT NON_TER 110
YT SEQUENCE 110 AA; 12652 MW; 7EEOA7E971D99552 CRC64;

Query Match 100.0%; Score 41; DB 12; Length 110;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 ELRRKMY 8
Y 53 ELRRKMY 60

RESULT 14
Q9YRMO;
ID Q9YRMO; PRELIMINARY; PRT; 111 AA.
AC Q9YRMO;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE Major immediate-early protein (Fragment).
DE Human cytomegalovirus.

```

```
C Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
C Betaherpesvirinae; Cytomegalovirus.
X NCBI_TaxID=10359;
N [1]
P SEQUENCE FROM N.A.
C STRAIN=241/94, 1366/90, and 185/85;
C Zweygberg Wirtgart B., Brytting M., Linde A., Wahren B., Grillner L.;
T "Sequence variation within three important cytomegalovirus gene regions
T in isolates from four different patient populations.";
L Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
R EMBL; AF099593; AAD04515.1; -.
R EMBL; AF099564; AAD04486.1; -.
R EMBL; AF099566; AAD04488.1; -.
T NON TER 1
T NON TER 111
Q SEQUENCE 111 AA; 12770 MW; 09E95018A5108566 CRC64;

Query Match 100.0%; Score 41; DB 12; Length 111;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 ELRRKMY 8
b 51 ELRRKMY 58

RESULT 15
9YRK6 PRELIMINARY; PRT; 111 AA.
C QYRK6
T 01-MAY-1999 (TrEMBLrel. 10, Created)
T 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
T 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
E Major immediate-early protein (Fragment).
S Human cytomegalovirus.
C Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
C Betaherpesvirinae; Cytomegalovirus.
X NCBI_TaxID=10359;
N [1]
P SEQUENCE FROM N.A.
C STRAIN=131/91;
C MEDLINE=99036778; PubMed=9817892;
A Zweygberg Wirtgart B., Brytting M., Linde A., Wahren B., Grillner L.;
T "Sequence variation within three important cytomegalovirus gene
T regions in isolates from four different patient populations.";
L J. Clin. Microbiol. 36:3662-3669(1998).
R EMBL; AF099583; AAD04505.1; -.
R GO; GO:0016491; F:oxidoreductase activity; IEA.
R GO; GO:0008152; P:metabolism; IEA.
R InterPro; IPR002198; ADH short.
R PROSITE; PS00061; ADH_SHORT; 1.
T NON TER 1
T NON TER 111
Q SEQUENCE 111 AA; 12855 MW; 789C9F858D69A523 CRC64;

Query Match 100.0%; Score 41; DB 12; Length 111;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 ELRRKMY 8
b 51 ELRRKMY 58
```

earch completed: May 21, 2004, 12:49:27
ob time : 26.3913 secs